MCTDH on-the-fly: Efficient grid-based quantum dynamics without pre-computed potential energy surfaces

Gareth W. Richings[1, a)] and Scott Habershon[1, b)]
Department of Chemistry and Centre for Scientific Computing, University of Warwick, Coventry, CV4 7AL, UK

We present significant algorithmic improvements to a recently-proposed direct quantum dynamics method, based upon combining well established grid-based quantum dynamics approaches and expansions of the potential energy operator in terms of a weighted sum of Gaussian functions. Specifically, using a sum of low-dimensional Gaussian functions to represent the potential energy surface (PES), combined with a secondary fitting of the PES using singular value decomposition, we show how standard grid-based quantum dynamics methods can be dramatically accelerated without loss of accuracy. This is demonstrated by on-the-fly simulations (using both standard grid-based methods and MCTDH) of both proton transfer on the electronic ground state of salicylaldimine and the non-adiabatic dynamics of pyrazine.

I. INTRODUCTION

Many chemical processes require a quantum-mechanical description to provide the full picture of the molecular dynamics; examples include proton tunnelling in enzyme catalysis[1,2], photochemistry of DNA molecules[3,4] or plant sunscreens[5,6], and other aspects of femtochemistry[7–9]. The most direct route to performing quantum dynamical simulations is to provide solutions to the nuclear, time-dependent Schrödinger equation (TDSE)

$$i\hbar\frac{\partial\Psi\left(\mathbf{q},t\right)}{\partial t}=\hat{H}\Psi\left(\mathbf{q},t\right), \tag{1}$$

where $\mathbf{q}$ is a vector representing the nuclear degrees-of-freedom (DOFs) of the system in question. Once the wavefunction, $\Psi(\mathbf{q},t)$ is known then one can extract any information required about the time-evolution of the system.

Solution of the TDSE using so-called grid-based methods is now quite standard; here, the wavefunction is expanded and evolved on a grid of localised basis functions, regularly distributed through configuration space[10]. The multi-configuration time-dependent Hartree (MCTDH) method,[10–14] a natural extension of grid-based methods which simultaneously employs time-dependent basis functions, has become the *de facto* gold standard in the area of wavefunction dynamics. MCTDH is capable of treating a few tens of DOFs and converges to the numerically exact solution of the TDSE (given a suitable basis of sufficient size). MCTDH has been used to study, for example: the photochemistry of pyrazine[15–17], pyrrole[18] and benzene[19,20]. The upper limit on the system size treatable by MCTDH has been extended by the introduction of the multi-layer MCTDH method[21–24], bringing systems of several hundred DOFs into the purview of fully quantum simulations.

However, there remain significant barriers to using grid-based quantum dynamics methods which mean that their more general uptake by wider computational and experimental communities remains low. The key challenge of using grid-based quantum dynamics is the requirement that the system Hamiltonian must be fully defined before performing wavepacket propagation. First, only a subset of the DOFs of the molecular system are studied to make the dynamics calculation tractable, so one is faced with a choice as to which DOFs might be important; this is an important challenge when encountering a previously-unstudied system. Second, because quantum mechanics is non-local, the global form of the PES must be known prior to performing the dynamics. In simulations using grid-based methods such as MCTDH, the global PES is usually determined by fitting pre-defined functions to a set of electronic energies calculated at various molecular geometries, for example by using the vibronic-coupling Hamiltonian (VCHAM) method which is based on PES expansion in low-order polynomials[18,25–27]. This fitting of the PES is an extremely arduous task which must be performed before one can even begin to perform MCTDH simulations; this is the central barrier to making grid-based methods like MCTDH more widely applicable. As an aside, it is worth noting that a similar difficulty can arise for the kinetic energy operator, although this can often be addressed by an appropriate choice of (*e.g.* rectilinear) coordinates before commencing PES fitting.[28]

To make accurate, grid-based quantum dynamics methods (*e.g.* MCTDH) more accessible and more widely applicable, our recent efforts have focused on coupling aspects of direct-dynamics (DD) simulations with standard-grid-based approaches. In DD methods, the PES is calculated "on-the-fly" during propagation of the dynamical wavefunction;[29,30] examples include trajectory surface hopping (TSH[31–42]), *ab initio* multiple spawning (AIMS[43–46]), the DD-variational multi-configuration Gaussian (DD-vMCG) method[47–58], and related Gaussian wavepacket-based approaches. These approaches all have the advantage of only requiring PES evaluations where and when they are needed during

a) Electronic mail: G.Richings@warwick.ac.uk
b) Electronic mail: S.Habershon@warwick.ac.uk

wavefunction propagation. However, these DD methods also have features which can limit their domain of application; for example, TSH and AIMS rely on classical equations-of-motion (EOMs) to propagate trajectories or basis functions, such that tunnelling and zero-point energy conservation cannot be treated explicitly, while the (fully quantum-mechanical) DD-vMCG method can suffer from numerical instability and linear dependence in the solution of the EOMs due to the non-orthogonality of the Gaussian basis functions.[49]

In contrast, our recently-proposed methodology combines the stability of grid-based quantum dynamics, particularly MCTDH, with the convenience of DD methods[59,60]. Our method relies on the ideas of Gaussian process regression (GPR) and kernel ridge regression (KRR)[61–68] to express the PES as a weighted sum of Gaussian functions; we refer to our overall approach hereafter as the DD-grid-based (DD-GB) method (with specific names, defined later, used when employing a particular quantum propagation method). Most importantly, the Gaussian-kernel-based representation of the PES used in our simulation strategy is capable of constructing a global PES using just a few local PES evaluations, and is therefore employable as a DD strategy; in addition, the results Gaussian-based representation of the PES is already in the sum-of-products form which is required for efficient MCTDH propagation. Furthermore, as discussed below, this DD-GB strategy is also compatible with on-the-fly strategies for dealing with non-adiabatic transitions, and so may be used to simulate photochemical dynamics as well as ground-state dynamics.

Our DD-GB method has already been shown to give accurate dynamical results when compared to benchmark studies using both fitted and *ab initio* on-the-fly PESs; however, our initial investigations also noted that the existing approach has significant computational demands associated with PES construction and wavefunction propagation. The aim of this work is to demonstrate two new algorithmic improvements to our DD-GB method which dramatically decrease computational effort, and hence greatly increase the potential utility of our approach. First, in Section III A, we show how additive kernels can be exploited in our KRR PES representation to reduce the number and dimensionality of integrals required for wavefunction propagation on a grid. Second, in Section IV, we show how singular value decomposition (SVD) can be used to generate a compact representation of KRR PESs generated on-the-fly during wavefunction propagation. Together, these two extensions are demonstrated to yield versions of GB and MCTDH propagations which are both accurate and essentially DD in nature; this is highlighted in DD-GB simulations of proton transfer in salicylaldimine and non-adiabatic dynamics of pyrazine.

II. METHODOLOGY

A. Grid-Based Quantum Dynamics: The Standard Method

The standard grid-based method for quantum dynamics is well established,[10] and we give here only the details relevant to the simulations presented below. For a molecular system with f DOFs, the wavefunction is expanded in terms of products of *time-independent* basis functions, each of which has an associated, time-dependent, complex coefficient, $C_{j_1,\cdots,j_f}$. For a nuclear wavepacket moving on electronic state, s, we then have

$$\begin{aligned}\Psi^{(s)}(q_1,\cdots,q_f,t) &= \sum_{j_1}^{N_1}\cdots\sum_{j_f}^{N_f} C^{(s)}_{j_1,\cdots,j_f}(t)\prod_{\kappa=1}^{f}\chi^{(\kappa)}_{j_\kappa}(q_\kappa)\\ &= \sum_{J} C^{(s)}_{J}(t)\,X_J(\mathbf{q})\end{aligned} \tag{2}$$

Note that we have taken the opportunity to introduce a compound index, $J{=}j_1,\cdots,j_f$. The total wavefunction for a system with N_s orthonormal electronic states is defined as

$$|\Psi\rangle = \sum_{s=1}^{N_s}|\Psi^{(s)}\rangle|s\rangle \tag{3}$$

Writing the total Hamiltonian as

$$\hat{H} = \sum_{su}^{N_s}|s\rangle\hat{H}^{(su)}\langle u|, \tag{4}$$

and employing the Dirac-Frenkel variational principle (DFVP),[69,70] we get a set of coupled EOMs for the expansion coefficients:

$$i\hbar\dot{C}^{(s)}_J = \sum_{u}^{N_s}\sum_{L}\langle X_J|\hat{H}^{(su)}|X_L\rangle C^{(u)}_L \tag{5}$$

Integration of these EOMs allows one to follow the propagation of the wavefunction through time. Given an appropriate basis of sufficient size, the propagation of these EOMs provides a numerically exact solution of the TDSE for a given Hamiltonian, and, as such, this method is referred to as the standard method (SM)[10].

B. Grid-Based Quantum Dynamics: MCTDH

Because of the exponential scaling with system size, the SM is restricted to treating about 5 DOFs in practical calculations. To enable simulations of larger molecular systems, MCTDH was introduced. In MCTDH, as in the SM, the wavefunction is expanded in a sum of products of

basis functions, each product having a complex expansion coefficient,

$$\begin{aligned}\Psi^{(s)}(Q_1,\cdots,Q_f,t)&\\ =\sum_{j_1}^{n_1}\cdots\sum_{j_m}^{n_m}A^{(s)}_{j_1,\cdots,j_m}(t)\prod_{\kappa=1}^{m}\varphi^{(s,\kappa)}_{j_\kappa}(Q_\kappa,t)&\\ =\sum_J A^{(s)}_J(t)\,\Phi^{(s)}_J(\mathbf{Q},t)\,.&\end{aligned}\tag{6}$$

The key difference between the SM and MCTDH *ansätze* is that the basis functions in the MCTDH method are *time-dependent*. Combining the MCTDH *ansatz* and DFVP yields EOMs for the coefficients and the time-dependent basis functions (referred to as single-particle functions (SPFs)):

$$i\hbar\dot{A}^{(s)}_J=\sum_u^{N_s}\sum_L\langle\Phi^{(s)}_J|\hat{H}^{(su)}|\Phi^{(u)}_L\rangle A^{(u)}_L\tag{7a}$$

$$i\hbar\dot{\boldsymbol{\varphi}}^{(s,\kappa)}=\left(1-P^{(s,\kappa)}\right)\left(\boldsymbol{\rho}^{(s,\kappa)}\right)^{-1}\sum_u^{N_s}\langle\hat{H}^{(su)}\rangle^{(\kappa)}\boldsymbol{\varphi}^{(u,\kappa)}\tag{7b}$$

The SPFs $\boldsymbol{\varphi}^{(s,\kappa)}$ are functions of a small subset (usually 1-4) of the system DOFs, $Q_\kappa=(q_{\kappa_1},\cdots,q_{\kappa_p})$, and are themselves expansions in terms of a time-independent basis (as in the SM wavefunction in Eq. (2)):

$$\varphi^{(s,\kappa)}_{j_\kappa}(Q_\kappa,t)=\sum_{i_\kappa}^{N_\kappa}c^{(s,\kappa,j_\kappa)}_{i_\kappa}(t)\,X^{(\kappa)}_{i_\kappa}(Q_\kappa)\,.\tag{8}$$

In the SPF EOMs (Eq. (7b)), $\hat{P}^{(s,\kappa)}$ is a projection operator onto the SPF space along mode κ, and $\left(\boldsymbol{\rho}^{(s,\kappa)}\right)^{-1}$ is the inverse of the density matrix associated with κ. By constructing a Hartree product of SPFs in all modes apart from κ, $\Phi^{(s)}_{J^\kappa}$, we can define a set of single-hole functions, $\Psi^{(s,\kappa)}_l=\sum_{J^\kappa}A^{(s)}_{J^\kappa_l}\Phi^{(s)}_{J^\kappa}$, from which a mean-field matrix, with elements $\langle\hat{H}^{(st)}\rangle^{(\kappa)}_{jl}=\langle\Psi^{(s,\kappa)}_j|\hat{H}^{(st)}|\Psi^{(t,\kappa)}_l\rangle$, is defined.

By evolving the SPFs variationally, it is possible to keep their number to a minimum; it is this reduction in the size of the basis in Eq. (6) which permits the study of larger systems than could be treated by the SM (MCTDH also scales exponentially, but with a smaller base than the SM[10]).

C. PESs for Grid-Based Wavefunction Propagation

To solve the EOMs in Eqs. (5) and (7) as efficiently as possible, the Hamiltonian must be in a sum-of-products form (*i.e.* all terms are products of functions of single DOFs), so that the multi-dimensional integrals can be reduced to sums-of-products of one-dimensional integrals. This feature of the Hamiltonian is not usually a problem for the kinetic energy part if a sensible choice of coordinate system is made, but this requirement can be difficult to ensure for the potential energy part of the Hamiltonian.

In recent work[59,60] we have presented a grid-based DD method using Eqs. (5) and (7) where the PES is constructed on-the-fly by using KRR fit to a set of PES values calculated at appropriately chosen molecular geometries, $\{\mathbf{q}^k\}$, at which the Gaussian kernel functions of the PES are also centered. Defining a one-dimensional kernel function along DOF, κ, as

$$k(q_\kappa,q^k_\kappa)=\exp(-\alpha(q_\kappa-q^k_\kappa)^2)\tag{9}$$

the potential energy operator is then represented as

$$\begin{aligned}V^{(su)}(\mathbf{q})&\approx\sum_{k=1}^{M}w^{(su)}_k\prod_{\kappa=1}^{f}k(q_\kappa,q^k_\kappa)\\ &=\sum_{k=1}^{M}w^{(su)}_k k(\mathbf{q},\mathbf{q}^k)\end{aligned}\tag{10}$$

The width parameter, α, can in principle be optimized by log-likelihood maximization, as employed in GPR, although we have found to date that using an appropriate fixed value of α, chosen at the outset, is sufficient for our purposes. The weights of the expansion, $\{w^{(su)}_k\}$, are determined by solution of the linear equation

$$\mathbf{K}\mathbf{w}=\mathbf{b},\tag{11}$$

where the vector, $\mathbf{w}$, contains the weights and the elements of the covariance matrix, $\mathbf{K}$, are[62]

$$K_{ij}=k(\mathbf{q}^i,\mathbf{q}^j)+\gamma^2\delta_{ij},\tag{12}$$

with γ, being a small, positive regularisation parameter. The vector, $\mathbf{b}$, contains the PES values calculated (*e.g.* using *ab initio* electronic structure calculations) at sampled geometries,

$$b_i=V^{(su)}\left(\mathbf{q}^i\right).\tag{13}$$

Returning to Eq. (10), we find that, if we use rectilinear coordinates (such as Cartesian coordinates or normal mode coordinates) in the dynamics calculation, the representation of the potential energy operator has the sum-of-products form necessary for efficient integral evaluation,

$$\langle X_J|\hat{V}^{(su)}|X_L\rangle=\sum_{k=1}^{M}w^{(su)}_k\prod_{\kappa=1}^{f}\int\mathrm{d}q_\kappa\,\chi^*_{j_\kappa}k(q_\kappa,q^k_\kappa)\chi_{l_\kappa}.\tag{14}$$

The individual integrals can be further simplified by an appropriate choice of time-independent basis; the implementation of the SM and MCTDH in the Quantics code[71] uses discrete variable representation (DVR) basis

functions[10], which are centered at points in configuration space, $\mathbf{q}_\alpha$ yielding a grid along each DOF, such that

$$\langle\chi_\alpha|\hat{V}^{(su)}|\chi_\beta\rangle \approx V\left(\mathbf{q}_\alpha\right)\delta_{\alpha\beta}. \tag{15}$$

This property means that the integrals in Eq. (14) can be evaluated simply by using the value of each term in the potential energy function at the location of each DVR gridpoint.

The KRR-based approach outlined above requires one to select the molecular geometries used in the PES fit. In DD-GB, this is implemented by randomly sampling a set number of geometries around the centre of the wavepacket at pre-determined intervals, and then using those geometries to calculate PES reference values. However, to reduce computational effort through calculating excessive numbers of potential energies, it is possible to avoid placing new reference points in regions of configuration space where the PES representation is already sufficiently accurate, we first compute the variance function[62]

$$\sigma^2\left(\mathbf{q}\right) = k(\mathbf{q},\mathbf{q}) + \gamma^2 - \mathbf{k}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{k}. \tag{16}$$

at the randomly-sampled geometries, $\mathbf{q}$. Here, γ and $\mathbf{K}$ are defined as in Eq. (12) whilst the vector, $\mathbf{k}$, contains the covariances of the new point with all of the pre-existing points (Eq. (9)). To avoid unnecessary PES evaluations in our scheme, we define a numerical parameter which, if exceeded by the variance at $\mathbf{q}$, signals that a new energy is required to be calculated; otherwise, the selected geometry is rejected and another sampled.

The key feature of this KRR-based methodology is that it brings GB/MCTDH simulations closer to the DD approach. One does not need to pre-fit or pre-compute a global PES; instead, the potential energy operator is learnt on-the-fly using PES evaluations (usually *ab initio* calculations) at configurations sampled from the evolving wavepacket. As a final point, we note that, after an initial wavepacket propagation and PES construction simulation has been performed, one can restart the simulation from the initial wavepacket using the full KRR-PES to obtain a final picture of the quantum dynamics.

1. Non-Adiabatic Simulations

To study multiple electronic states, and the non-adiabatic transitions between them, using grid-based methods, it is desirable to use a quasi-diabatic PES where the discontinuities in the gradient of the adiabatic surfaces and the non-adiabatic couplings are transformed away, yielding smoothly crossing surfaces. To achieve such a transformation on-the-fly, we use a modified version of a scheme proposed by one of us[72,73] in the context of DD-vMCG simulations. This approach is based on propagation of the diabatization matrix,[74] $\mathbf{A}$, using line integrals of the non-adiabatic coupling terms between the adiabatic states, ψ_i and ψ_j (with respective energies V_{ii}^A and V_{jj}^A), given by

$$\mathbf{F}_{ij} = \frac{\langle\psi_i|\nabla\hat{H}|\psi_j\rangle}{V_{jj}-V_{ii}}. \tag{17}$$

The approximate relationship[75]

$$\nabla\mathbf{A} \approx -\underline{\mathbf{F}}\mathbf{A}, \tag{18}$$

holds for an incomplete set of adiabatic states (it is an equality for a complete set) where the underlining of $\underline{\mathbf{F}}$ indicates that it is a matrix of vectors. As we are interested in non-radiative transfers between a small number of (usually two) energetically-close states, we take the pragmatic view of using Eq. (18) as an equality in the method here. Starting at a geometry, $\mathbf{q}$, where $\mathbf{A}$ is known, Eq. (18) is integrated between $\mathbf{q}$ and some new point, $\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}$, to give $\mathbf{A}$ at the new point:

$$\mathbf{A}\left(\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}\right) = \exp\left(-\int_{\mathbf{q}}^{\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}}\underline{\mathbf{F}}\cdot\mathrm{d}\mathbf{q}\right)\mathbf{A}\left(\mathbf{q}\right). \tag{19}$$

The matrix, $\mathbf{A}$, then allows transformation of the adiabatic energy matrix, $\mathbf{V}^{\mathbf{A}}$, to the quasi-diabatic representation

$$\mathbf{V}^{\mathrm{D}} = \mathbf{A}^{\mathrm{T}}\mathbf{V}^{\mathrm{A}}\mathbf{A}. \tag{20}$$

In practice the choice of $\mathbf{A}$=$\mathbf{I}$ is made at the center of the initial wavepacket and the transformation matrix propagated away from this point towards newly-sampled geometries as the evolution of the wavepacket and PES proceeds.

2. Problems with Computational Efficiency

Previous work[59,60] has already demonstrated that the scheme outlined above can reproduce the quantum dynamics results obtained by using pre-fitted PESs or DD-vMCG. However, we have also noted[60] the lack of efficiency of our original scheme, especially when coupled to MCTDH propagation of the wavefunction; even for a system comprising six DOFs and one electronic state (a modest system by normal MCTDH standards), we found that a propagation of just 100 fs could take several weeks to perform. Such time requirements clearly make the method, as described above, of limited applicability.

The inefficiency of our original DD-MCTDH scheme is caused by the large number of reference points needed to expand the PES (*i.e.* the large value of M in Eq. (14)). Assuming an MCTDH wavefunction constructed from one-dimensional SPFs (the arguments presented also follow for wavefunctions constructed using multi-dimensional SPFs), the PES contribution to the Hamiltonian integral in Eq. (7a), when using the KRR-based approach outlined above is then

$$\begin{aligned}&\langle\Phi_J^{(s)}|\hat{V}^{(su)}|\Phi_L^{(u)}\rangle \\ &\qquad= \sum_{k=1}^{M} w_k^{(su)}\prod_{\kappa=1}^{f}\int\mathrm{d}q_\kappa\varphi_{j_\kappa}^{(s,\kappa)*}k(q_\kappa,q_\kappa^k)\varphi_{l_\kappa}^{(u,\kappa)}\end{aligned} \tag{21}$$

Inserting Eq. (8) we get

$$\langle \Phi_J^{(s)} | \hat{V}^{(su)} | \Phi_L^{(u)} \rangle = \sum_{k=1}^{M} w_k^{(su)} \prod_{\kappa=1}^{f} \sum_{m_\kappa, n_\kappa}^{N_\kappa} c_{m_\kappa}^{(s,\kappa,j_\kappa)*} c_{n_\kappa}^{(u,\kappa,l_\kappa)} \times \int \mathrm{d}q_\kappa \chi_{m_\kappa}^{*} k(q_\kappa, q_\kappa^k) \chi_{n_\kappa} \tag{22}$$

Because the SPFs are time-dependent the transformations of the DVR integrals must be performed at *every* step of wavefunction propagation, an extremely time-consuming process. In an ideal world, the transformation to the SPF basis would be performed on the full DVR potential integral, in other words sum up the DVR integrals for all potential terms, then transform. To do so would require the two summations in Eq. (22) to swap place, or as the summation over m_κ, n_κ depends on the product index, at least for the GPR summation to be moved to the inner position. However, this is clearly not a valid change, so we are stuck with this time-consuming form. The problem is even worse when considering the mean-field matrices in Eq. (7b) because a large number of such matrices must be constructed in a similar way to Eq. (22).

Frustratingly, calculations using the SM do not suffer from this efficiency bottleneck because there is no need to perform a transformation to a time-dependent basis (see Eq. (5)). As described in the previous section, the potential energy operator is only updated occasionally during wavefunction propagation; once this has occurred the Hamiltonian integral in Eq. (5) can be evaluated, stored and used until the next update. This reflects the simplicity of the SM; it is only necessary to know the total value of the potential at each point on this product grid, not the separate contribution from each term, as required by MCTDH.

In summary, it is apparent that the best way to speed up DD-MCTDH calculations using the KRR expansion of the PES is to reduce the number of terms needed (obviously, without losing accuracy). Such a reduction will have the added benefit that fewer *ab initio* electronic structure calculations need to be performed. In the next section, we present an approach for achieving the reduction in the size of the KRR database.

III. METHODOLOGY IMPROVEMENT I: ADDITIVE KERNELS

A. Theory and Implementation

As outlined above, initial implementation of the DD-GB method used f-dimensional Gaussian functions as the kernels to fit to the PES (Eq. (10) and (12)); we refer to this kernel hereafter as the *full* kernel. The large number of such functions required to represent a PES is due to the Gaussian function's locality; for a given value of α, the full-width at half-minimum (FWHM) of the Gaussian along any DOF is $2\sqrt{\alpha \ln 2}$. It follows that a reasonable measure of the region of influence of each Gaussian is given by the volume of the hypersphere of that radius,

$$\text{Volume} = \frac{\pi^{f/2}}{\Gamma(f/2+1)} 2^f (\alpha \ln 2)^{f/2}. \tag{23}$$

Given some representative length, l, of the DVR grids, the volume of configuration space in which the wavepacket propagation proceeds is thus l^f, meaning that each Gaussian kernel "occupies" a fraction of configurational space given by

$$\mathrm{P} = \frac{1}{\Gamma(f/2+1)} \left(\frac{2\,\pi^{1/2} (\alpha \ln 2)^{1/2}}{l} \right)^f \tag{24}$$

Because the Gaussian widths are less than the grid length along each DOF, it is clear that the proportion of configuration space influenced by each Gaussian kernel decreases exponentially with increasing number of DOFs. To reduce the number of Gaussian functions required to span configuration space, it is thus necessary to use a less localised kernel. Fortunately, the choice of an f-dimensional Gaussian kernel is just one of many possible within GPR and KRR[61], with the main restriction on the choice of kernel being that it must ensure the covariance matrix, $\mathbf{K}$, is symmetric and positive semi-definite. To this end, it has been shown that a kernel constructed from a sum of lower dimensional Gaussian functions is a valid covariance function, as in

$$k^{\mathrm{Add}}(\mathbf{q}, \mathbf{q}^n) = \sum_{\kappa=1}^{f} k(q_\kappa, q_\kappa^n) + \sum_{\kappa,\lambda}^{f} k(q_\kappa, q_\kappa^n) k(q_\lambda, q_\lambda^n) + \cdots + k(\mathbf{q}, \mathbf{q}^n). \tag{25}$$

This type of expansion, termed an additive kernel, is reminiscent of that used within the high-dimensional model representation (HDMR)[76–78], a fit being built up from a set of increasingly higher-dimensional functions. Fitting a PES to a set of low-dimensional functions is also the approach used within the successful VCHAM approach, so such an approximation is valid when trying to fit multi-dimensional PESs, as well as fulfilling the requirement of being in the sum-of-products form for efficient MCTDH propagation. In practice, the additive kernel is simply used to replace the f-dimensional Gaussian functions in Eqs. (10) and (12); all other aspects of the KRR process remain as before.

How does such a kernel help in reducing the number of reference points required to construct a PES expansion? Returning to the volume of influence of the kernel function, consider one of the 1-dimensional terms, $\exp(-\alpha(q_\kappa - q_\kappa^n)^2)$. At some coordinate, q_κ^0, this term has the value $k(q_\kappa^0, q_\kappa^n)$, but as it is independent of the other $f-1$ coordinates it has the same value everywhere

within the configuration subspace defined as consisting of the points $(q_1, \cdots, q_\kappa^0, \cdots, q_f)$. The subspace has a volume of l^{f-1} therefore, based on the FWHM argument, the kernel has influence over a volume of $2\sqrt{\alpha \ln 2}\, l^{f-1}$; expressed as a proportion of the total configuration space, this volume is $2\sqrt{\alpha \ln 2}/l$. This volume is independent of the dimensionality of the problem at hand, so each reference point can describe a much larger proportion of the PES than if a single f-dimensional Gaussian was used. Similar arguments hold for all of the terms in Eq. (25) (except the final, f-dimensional term).

One can truncate the expansion in Eq. (25) at any order and still be left with a valid kernel; however, because two-body terms tend to dominate in expressions of the PES,[76] we propose here to use a two-body additive kernel (*i.e.* truncating the expansion in Eq. (25) after the second term) as an alternative to the f-dimensional Gaussian functions used previously. Extending our previous work,[59,60] we have implemented DD-SM and DD-MCTDH codes which employ a second-order additive kernel in a development version of the Quantics quantum dynamics package[71]. In the next section we present the results of calculations, using this implementation, which demonstrate the ability of the additive kernel to accurately represent the underlying PES with lower computational effort than the corresponding f-dimensional version employed in our previous work.

B. Salicylaldimine Proton Transfer

In this section, we present the results of wavefunction propagations performed on a 4-dimensional model of proton transfer in salicylaldimine using a development version of the Quantics quantum dynamics package[71]. Three different calculations were performed: (i) a reference SM calculation propagation using the VCHAM-fitted PES of Polyak *et al*[57], (ii) DD-SM on the same PES using a 4-dimensional kernel, and (iii) DD-SM on the same PES using a second-order additive kernel.

All calculations were performed in mass-frequency scaled normal mode coordinates; using the nomenclature of Polyak *et al* the in-plane modes ν_1 (proton transfer mode), ν_{13} (bending of O and N away from one another), ν_{32} (CO stretch and NH bend) and ν_{36} (OHN bending) were chosen as the system DOFs. A sine DVR of 101 gridpoints was used for the ν_1 mode whilst 21 member harmonic oscillator DVR bases were used for the other three modes. The initial wavepacket was a 4-dimensional Gaussian function centred at $\langle \nu_1 \rangle$=0.96 (to the enol side of the barrier), $\langle \nu_{13} \rangle = \langle \nu_{32} \rangle$=0 and $\langle \nu_{36} \rangle$=0.14, with widths of $\langle \mathrm{d}\nu_1 \rangle$=0.5706, $\langle \mathrm{d}\nu_{13} \rangle$=0.6902, $\langle \mathrm{d}\nu_{32} \rangle$=0.6707, $\langle \mathrm{d}\nu_{36} \rangle$=0.7704. Time-propagation was carried out for 100 fs; the DD-SM calculations sampled the PES every 1 fs, with 300 geometries being randomly selected within 3 times the width of the wavepacket from the centre in each DOF. PES values were calculated at the chosen geometries if the variance there was greater than 10^{-3}. A kernel width parameter, α, of 0.5 was used for all DD-SM calculations. Subsequently, DD-SM calculations were performed using the database of energies calculated previously, with the potential energy operator generated at the start of the calculation and with no further enlargement of the initial PES database. In all cases, the EOMs were integrated using the 15$^{\text{th}}$-order short iterative Lanczos (SIL) method with accuracy cutoff of 10^{-6}. A flux operator[10] was evaluated at q_1=0 during the propagation, allowing measurement of the flow of the wavefunction across the potential barrier for the proton transfer.

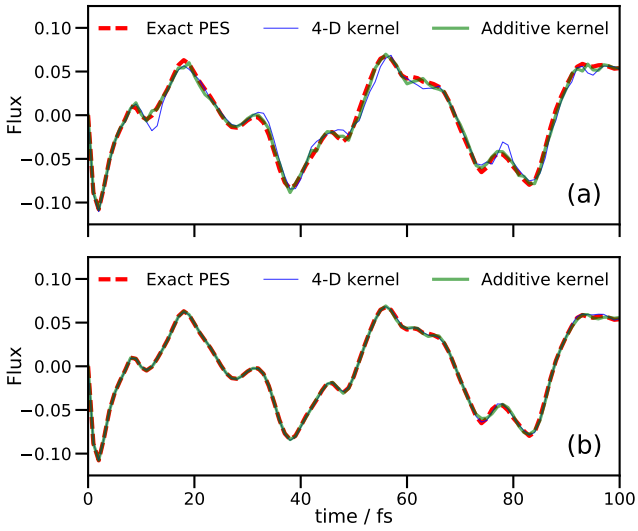

FIG. 1. Flux operator expectation value for a model of salicylaldimine with 4 DOFs. The dividing surface is placed at the barrier of the transition mode, ν_1. Both (a) and (b) show results of grid-based quantum dynamics using the standard grid method on the fitted PES (red dashed lines), a DD-SM simulation using a 4-dimensional KRR PES (thin blue line), and a DD-SM simulation using a second-order additive kernel (thick green line). In (a), DD-SM results are shown for the initial wavefunction propagation; in (b), DD-SM results are shown for a second simulation in which all reference points learned in (a) are used to construct the respective potential energy operators.

Figure 1(a) shows the flux operator expectation value along mode ν_1 for the full SM calculation and the two DD-SM calculations with different kernels; here, the energy database was built on-the-fly during the DD-SM calculations. All plots are qualitatively similar, with initial motion of the wavepacket across the barrier towards the keto tautomer, followed by the oscillation between the keto and enol sides of the barrier oscillates with a period of about 40 fs. The actual dynamics is as expected from previous work; the important point to note here is that the dynamics matches the reference SM results extremely well when using both the full kernel or the second-order additive kernel, with just small differences

appearing later on as inaccuracies due to the sampling of the potential creep in.

The key difference in performance between the kernels is in the calculation timings. Both DD-SM calculations employed a single core on a standard desktop machine; however, whilst the full-kernel calculation required 15.5 CPU hours, the additive kernel calculation took 2.25 CPU hours (an improvement by a factor of around 7). The origin of the difference in computational effort is the different number of points added to the database of electronic energies; the full kernel calculation generated 8,026 energies, but the additive kernel only required 741 geometries to be sampled to generate an accurate PES representation. This saving in the number of energy datapoints required in the two DD-GB methods accounts for nearly all of the difference in the computation time (in fact, the times required for wavefunction propagation were quite similar, 2,425 against 2,151 CPU seconds for the additive and full kernels, respectively). Furthermore, the calculation of the KRR variances and weights took nearly 10,000 CPU seconds with the full kernel, but only 36 CPU seconds with the additive kernel (this huge difference reflects the relative speeds of solving the linear equations in Eqs. (11) and (16) when using a matrix with 100 times fewer entries). Finally, the largest difference in timings was actually in calculating and storing the integrals in Eq. (15): 43,700 CPU seconds for the full kernel against 5,300 for the additive kernel. Both calculations used the same size DVR grid, so the difference in these times is solely due to the number of terms in the potential operator. In these calculations, where a pre-fitted, analytic potential is available, the individual energy calculations are trivial, but if this calculation were repeated using an *ab initio* electronic structure program, where each energy calculation can take anything from a few seconds to many minutes or more, the ability to reduce the number of calculations by an order of magnitude will result in a large saving of computational effort.

Figure 1(b) shows results for a second set of DD-SM simulations performed using the full databases generated by the respective full-kernel and additive kernel simulations from Fig. 1(a). The agreement between both calculations and the exact result is essentially perfect; the improvement of both fits compared to the original calculations is to be expected as a much larger region of configuration space has been used to generate the potential energy operator. In this secondary calculation, the time savings for the additive kernel over the full kernel are not as marked as in Fig. 1(a) (2,710 and 3,215 CPU seconds for full and additive kernels, respectively), because the computationally-demanding integral evaluation only occurs once, at the start of the propagation.

The conclusion of this section is that we find that the additive kernel is able to reproduce an underlying PES just as accurately as the full kernel, but requires far fewer energy evaluations; this result is very promising for grid-based quantum dynamics. However, initial attempts to combine the additive kernel method with MCTDH to allow DD-MCTDH simulations did not lead to any significant time-savings. In particular, when using the additive kernel, each database point gives rise to multiple, individual terms in the potential energy operator, one for each term in the summations in Eq. (25) (*e.g.* ten each in this case); these terms cannot be contracted down to a single operator when using MCTDH. As a result, when attempting to use the additive kernel in a DD-MCTDH scheme, the reduction in computational effort for the single-point energy evaluations and solution of the KRR linear equations is still achieved, but we do not gain anything in MCTDH propagation because there is little reduction in the number of terms in the potential energy operator.

In the next section we outline a further improvement which, in combination with the additive kernel, *is* able to reduce the number of terms in the potential energy operator, thereby allowing a significant speed up in the dynamics calculations, making DD-MCTDH calculations feasible.

IV. METHODOLOGY IMPROVEMENT II: SINGULAR VALUE DECOMPOSITION OF A GAUSSIAN PES

A. Theory and Implementation

Here, we present a new method of fitting a PES which provides another step in improving the efficiency of the DD-GB methods. Here, the KRR fitting of the PES is relegated to the secondary status of being an efficient sampling method which generates a representation of the PES; we then implement a second transformation from which another, more compact potential energy operator expression can be determined. In particular, we use the idea, demonstrated in Eq. (15), that, when using a DVR basis, we only need to know the value of the potential term at the location of the gridpoint.

To describe this second extension to our DD-GB approach, we assume that we have a KRR PES generated using the additive kernel approach described above. In the setting of our work, this KRR PES would be generated during a DD-GB simulation and the decomposition method described below can be used to accelerate evaluation of integrals over the potential energy operator (as required to accelerate MCTDH simulations). This KRR PES, referred to as $V^0(\mathbf{q})$, can then be decomposed into a simpler PES representation, appropriate for MCTDH propagation, using the following steps:

1. First, we evaluate the KRR potential energy function, V^{KRR}, at the origin of the coordinate system, $\mathbf{0}$, to set the relative, constant shift for the PES.

2. The KRR potential energy $V^{\mathrm{KRR}}(\mathbf{q})$ is evaluated at the location of the DVR gridpoints along *one* of the DOFs, with all other DOFs remaining at $q_f = 0$. This gives a one-dimensional potential energy term along

the selected mode (including any constant shift in the potential).

3. One-dimensional KRR PES slices along each of the remaining DOFs are then generated, but with the value of the KRR PES at the origin being subtracted, this time, to avoid over-counting. By this simple procedure we thus have a representation of the KRR PES as a sum of one-dimensional terms, with a single vector of numbers representing the PES along each DOF.

4. To obtain the 2-dimensional terms which couple wavepacket motion between the DOFs, we must fit the difference of the full KRR PES and the one-dimensional terms, between each unique pair of DOFs; in other words, we are fitting to the residual error between representation of the PES using only one-dimensional terms and the full KRR PES.

   To achieve this, consider two DOFs, q_g and q_h, with the coupling between them written as $V^{gh}(q_g, q_h)$. Assuming this function is separable, the integrals over the DVR basis, which appear in the wavefunction propagation, are:

$$\begin{aligned}\langle X_i^{(g)} X_j^{(h)} | V^{gh}(q_g, q_h) | X_k^{(g)} X_l^{(h)} \rangle \\ = \langle X_i^{(g)} | V^{gh}(q_g) | X_k^{(g)} \rangle \langle X_j^{(h)} | V^{gh}(q_h) | X_l^{(h)} \rangle \\ \approx V^{gh}(q_g^i) V^{gh}(q_h^j)\, \delta_{ik} \delta_{jl}\end{aligned} \tag{26}$$

   So, we find that these coupling terms, when evaluated in the DVR basis, are the product of terms on the DVR coordinate grid along the individual DOFs, g and h. We then define the residual function at any point (q_g^i, q_h^j) on this new, two-dimensional grid as

$$\begin{aligned}V^{gh}(q_g^i, q_h^j) = & V^{\mathrm{KRR}}(q_g^i, q_h^j) - V^g(q_g^i) - V^h(q_h^j) \\ & - V^{\mathrm{KRR}}(\mathbf{0})\end{aligned} \tag{27}$$

   If the DOFs, g and h, contain N_g and N_h DVR gridpoints, respectively, Eq. 27 defines an $N_g \times N_h$ matrix, $\mathbf{V}^{gh}$, the elements of which are the value of the residual at each point.

5. We can then decompose the matrix $\mathbf{V}^{gh}$ into a sum of outer products of two vectors, each term of which represents a contribution to the total coupling term. The vectors contain the values of each coupling term at the locations of the DVR gridpoints along the separate DOFs. To find such a set of vectors, we minimize the following squared Frobenius norm,

$$||\mathbf{V}^{gh} - \mathbf{V}_g^{gh} \otimes \mathbf{V}_h^{gh}||^2 \tag{28}$$

   by performing a singular value decomposition (SVD) of the coupling matrix

$$\mathbf{V}^{gh} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^T \tag{29}$$

   with the singular values in $\mathbf{\Sigma}$ being in non-decreasing order. It follows that

$$V^{gh}(q_g^i, q_h^j) = \sum_{k=1}^{\min(N_g, N_h)} \sigma_k\, u_{ik}\, w_{jk} \tag{30}$$

   We can thus set the elements of the required vectors, $\mathbf{V}_g^{gh}$ and $\mathbf{V}_h^{gh}$, which correspond to the values of the potential energy operator terms at the DVR gridpoints, to be

$$V_{g(k)}^{gh}(q_g^i) = \sqrt{\sigma_k} u_{ik} \tag{31a}$$

$$V_{h(k)}^{gh}(q_h^j) = \sqrt{\sigma_k} w_{jk} \tag{31b}$$

6. Finally, we note that, in practice, the term with the largest singular value is added, and then a new residual is created by subtracting this term from the original. If the norm of this new residual is below some pre-determined accuracy cutoff, no further terms are added; if the norm of the residual is larger than the cutoff, we loop over the other singular values in descending order, adding terms and checking the norm of the residual until convergence is reached.

7. The decomposition procedure above is repeated for all pairs of DOFs in the system.

The resulting PES representation is much more compatible with efficient MCTDH propagation because the multiple potential energy terms of the original KRR PES are contracted down to f one-dimensional terms and at most $\sum_{g<h}^{f} \min(N_g, N_h)$ two-dimensional terms. Finally, we note that this method has a similar philosophy to the POTFIT algorithm[10], which takes a PES evaluated on a grid and produces a PES in sum-of-product form by decomposing PES density matrices. An important similarity between the methods is that, in the context of grid-based dynamics methods such as MCTDH, the exact functional form of the potential is less important than obtaining the values of the potential terms at the locations of the DVR gridpoints. Of course, the underlying KRR PES used in our DD strategies is already in sum-of-products form, albeit with a sum which has too many terms for efficient simulation.

B. Salicylaldimine Proton Transfer: 4D

In this section, we demonstrate the improved computational efficiency of the SVD PES fitting method over the standard KRR approach, whilst simultaneously showing that the new method does not lead to a reduction in accuracy in the quantum dynamics.

As our benchmark problem, we again consider the 4-dimensional model of salicylaldimine proton transfer, as discussed in Section III B. We use the same computational setup with regards to DVR grid, choice of DOFs,

initial wavefunction and GPR sampling, although we focus here on the second-order additive kernel only. In addition, we also us the same fitted PES as that considered in Section III B. Wavefunction propagation was performed using MCTDH with two sets of 14 SPFs each; one describing modes ν_1 and ν_{36}, the other ν_{13} and ν_{32}. The variable mean field[10] implementation of the 6$^{\text{th}}$-order Adams-Bashforth-Moulton (ABM) integrator was used to solve the EOMs with an accuracy parameter of 10^{-5} and initial step length of 10^{-4} fs. A database of energies was created on-the-fly during propagation, and the SVD procedure, described above, was used to generate the potential energy operator for the dynamics from the KRR fit using the additive kernel. Terms from the SVD fit (Eq. (30)) were added until the Frobenius norm of the residual over all gridpoints was less than 10^{-3}. Following an initial pass in which the PES was generated on-the-fly, a second propagation was then performed using the full database of energies generated previously, with the PES fit performed at the start of the propagation and not updated further.

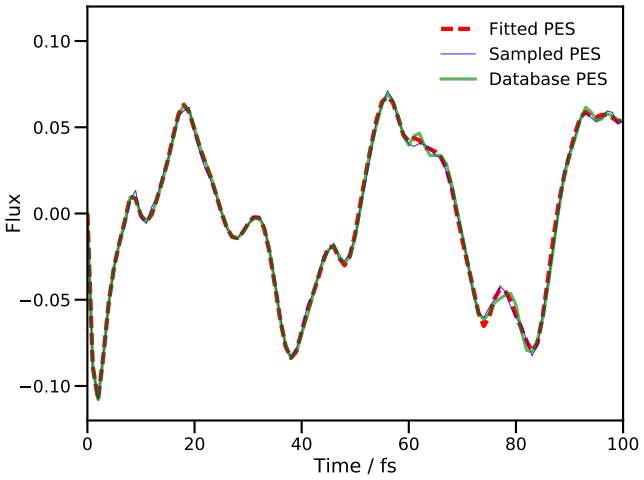

FIG. 2. Calculated proton transfer flux operator expectation value for a four DOF model of salicylaldimine. The dividing surface is placed at the barrier of the transition mode, ν_1. The dashed, red line is the result of using standard MCTDH with the fitted PES. The thin blue line is the result of a DD-MCTDH calculation with the potential energy operator constructed using SVD/additive kernel, and the thick green line uses the same SVD method, but using the PES database generated in the previous sampling calculation. All simulations used energies evaluated using the fitted PES of Polyak *et al.*[57]

In Fig. 2 we present the calculated flux expectation value along the ν_1 mode for both calculations, as well as the result from an MCTDH calculation using the same PES. There is little to say about the results in Fig. 2, except that the DD-MCTDH calculations give results in excellent agreement with the standard MCTDH calculation.

With regards to computational effort, 741 energy values were added to the PES database during the first DD-MCTDH propagation. This calculation took 3,893 CPU seconds on a single-core desktop computer, with the SVD fitting routines contributing only 59 CPU seconds to the total; in other words, the additional effort from SVD fitting is minimal. The next question is whether there is a saving in effort given by reducing the number of terms in the potential energy operator. For a one-state, four-DOF problem, there are ten potential terms for each energy value added to the database when using the standard KRR fit with the second-order, additive kernel; in the calculation reported here, this means the potential energy operator without SVD reduction would comprise 7,410 terms.

During the first calculation, when the PES is initially sampled, the number of potential energy operator terms, generated by the new SVD procedure, ranged from 39 to 56 (with a different number generated after each sampling step). The second calculation, using the pre-computed database, used a fit of 56 terms, a reduction by a factor of more than 130 over the standard KRR fit. It is also worth noting that the VCHAM fitted PES, from which we extracted the energies used in these calculations, had 40 potential operator terms (19 1-dimensional terms and 21 2-dimensional, compared to a split of 4 and 52 in the SVD fit). The SVD fit in this case requires only a few more terms than the VCHAM fit and, in fact, the DD-MCTDH calculation with the pre-computed database took less time than the standard MCTDH calculation (1,222 against 2,098 CPU seconds) due to shorter integration time-steps being taken in the latter case.

C. Salicylaldimine Proton Transfer: 6D

Having demonstrated the accuracy and reduced computational effort, we provide a further example to compare with results from our previous work[60]. In this previous case, we considered a 6-dimensional model of proton transfer in salicylaldimine, performing MCTDH dynamics using the fitted PES, and the standard KRR fit with full kernel. In that work, we found reasonable agreement between the full, MCTDH dynamics and the DD-MCTDH calculation, but with considerable computational effort. As such, we repeat these calculations here using our new DD-MCTDH scheme to further assess the benefits.

The 6-dimensional calculations included the same DOFs as used in the 4-dimensional calculations discussed above, but with addition of modes ν_{10} and ν_{11} (see Fig. (2) in reference 57). These additional modes each required 21 basis functions (harmonic oscillator DVR) and were included in an additional 2-dimensional combined mode described by 14 SPFs. The four original DOFs were treated in the same way as described in Sections III B and IV B, as was the integration of EOMs. The ini-

tial wavepacket was the product of Gaussian functions in all six DOFs, with centres and widths as described above for the four original DOFS and, additionally centred at $\langle v_{10}\rangle$=$\langle v_{11}\rangle$=0.0, with additional widths $\langle \mathrm{d}v_{10}\rangle$=0.7745 and $\langle \mathrm{d}v_{11}\rangle$=0.7590.

Figure 3 shows the flux along v_1 for three calculations: (i) standard MCTDH using the fitted PES, (ii) DD-MCTDH using the additive kernel and SVD fitting, with a database of energies from a prior calculation, and (iii) DD-MCTDH using the standard, 6-dimensional kernel with a pre-computed database of energies and using 18 SPFs per mode, as presented previously in reference 60.

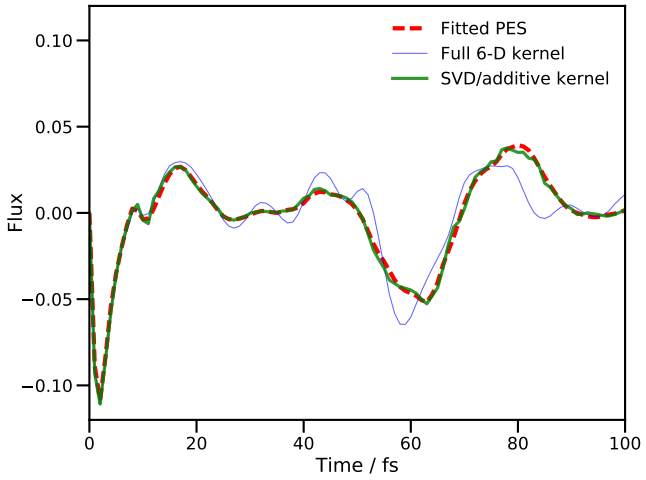

FIG. 3. Plot of the flux operator expectation value for a six degree-of-freedom model of salicylaldimine. The dividing surface is placed at the barrier of the transition mode, v_1. The solid, red line is the result of using MCTDH with the fitted PES. The hashed, blue line is the result of a DD-MCTDH calculation with the potential operator constructed using the SVD fit to an additive kernel GPR PES and the dashed, green line is the equivalent result running the calculation with the pre-calculated database. All calculations used energies calculated from the fitted PES of Polyak *et al*[57].

Figure 3 clearly demonstrates the improvement in agreement between the full MCTDH result and the DD-MCTDH calculations when using the SVD fit and additive kernel; the RMS errors in the calculated fluxes are an order of magnitude smaller when using SVD and additive kernel, compared to the full kernel calculation. We note that the effect of using 14 SPFs per combined mode as opposed to 18 in the full kernel DD-MCTDH calculation is minimal; an 18 SPF MCTDH calculation was also carried out and the plot of the flux was found to be almost identical to the corresponding 14 SPF calculation.

The errors in the full kernel case are due to the lack of convergence in the database of energies[60]; 10,101 entries were sampled, the maximum possible in that calculation, and the propagation took 143 hours on 16 processors. As a result, achieving convergence of the database would require a huge additional computational effort. In contrast, the SVD/additive kernel calculation only needed a database of 2,185 energies, once again illustrating the large saving in effort by using an additive kernel. From these 2,185 energies, a potential expansion of 134 terms was constructed by the SVD procedure, comparing very favourably with the 45,885 needed when using the additive kernel alone (six 1-dimensional, and fifteen 2-dimensional, terms *per* database entry). We also note that the VCHAM-fitted PES contained 78 terms in the six DOFs (29 1-dimensional and 49 2-dimensional), so the DD-MCTDH procedure does require more terms in the PES expansion, leading to a longer propagation (4,451 CPU seconds compared to 3,258 in the standard MCTDH calculation, both on a single processor on a desktop machine). Just over 12 hours CPU time was required for the initial propagation which generated the database for PES construction in the DD-MCTDH calculation, meaning that the full DD-MCTDH procedure with SVD and additive kernel took about 13 and a half hours; this is far less computational effort than required to perform either VCHAM fitting or a DD-MCTDH calculation using the full kernel. We also note that the use of 14 SPFs *per* combined mode as opposed to 18 for the full kernel calculation produces a saving in effort, all things being equal, but only by a factor of about 2.7 (based upon Eq. (74) in reference 10); nowhere near enough to account for the actual difference, which is mainly down to the size of the database. The use of SVD fitting with additive kernel in DD-MCTDH thus improves accuracy while simultaneously reducing computational effort.

D. Non-Adiabatic Dynamics of Pyrazine

As a final example showing the utility of the proposed SVD/additive kernel variant of the DD-MCTDH approach, we present results of a simulation modelling the non-adiabatic dynamics of pyrazine using *ab initio* electronic structure calculations. Pyrazine is a classic test case in non-adiabatic dynamics, particularly in the calculation of the absorption spectrum obtained by excitation to the S_2 excited electronic state.[15–17]. The presence of a conical intersection with the S_1 state in the vicinity of the Franck-Condon geometry makes this system an ideal test of whether a method can accurately represent the wavepacket as it moves between the electronic states.

To test our method, the geometry of pyrazine was optimized at the state-averaged complete active space self-consistent field level with 8 electrons in 8 orbitals (SA-CAS(8,8)) using the DZP basis in Molpro[79,80], the three lowest energy states being included and equally weighted. The orbitals were those in the π-system, plus the lone pair orbitals on the nitrogens. A subsequent frequency calculation generated the normal modes for use in the dynamics. A DD-MCTDH calculation was then performed

on the two excited states using the four mass-frequency scaled normal-modes v_{6_a}, v_{10_a}, v_1 and v_{9_a} (given as modes $3A_g$, $7B_{1g}$, $10A_g$ and $15A_g$ by Molpro), as used in the original MCTDH studies. Harmonic oscillator DVR basis sets were used along all four modes with 32, 22, 21 and 12 functions, respectively, One-dimensional SPFs were used for all DOFs, with different functions on each state (the multi-set formalism[10]): 7 SPFs on each state were used along v_{6_a}, with 12 on S_1 and 11 on S_2 on v_{10_a}, 6 and 5 respectively along v_1, and sets of 5 and 4 members each on v_{9_a}. The default ABM integrator was used to solve the MCTDH EOMs for 100 fs, with the initial wavefunction constructed as a product of the v=0 harmonic oscillator eigenfunctions along each mode, placed on the S_2 state and centered at the Franck-Condon point. To generate energies to construct the KRR representation of the PES, 100 points within 3 standard deviations of the wavepacket centre were sampled every femtosecond and added to the database if the variance at that geometry exceeded 10^{-3}. The energies of the two electronic states, along with the non-adiabatic couplings between them, were calculated using SA-CAS(8,8)/DZP, as described above; the energies were diabatised as outlined in Section II A prior to performing KRR fitting. Additionally, in order to maintain symmetry in the PES, each geometry was reflected in the $v_{6_a}/v_1/v_{9_a}$ plane (*i.e.* the sign of the v_{10_a} coordinate was changed and a point added at the resulting geometry). The additive kernel was used in the KRR process, and the SVD fitting procedure was used to generate the final PES on which the dynamics were performed.

As above, two calculations were performed; the first to generate the database of energies needed to represent the PES, and the second using the full database. Overall, 5,185 energies were generated during the first propagation, resulting in 335 terms in the PES expansion generated using SVD. Results from both calculations are shown in Figs. 4 and 5.

In Fig. 4, we show the absorption spectrum after vertical excitation to the S_2 state, calculated in both calculations using the Fourier transform of the autocorrelation function

$$c(t) = \langle \Psi(0) | \Psi(t) \rangle. \tag{32}$$

More details of this method can be found in reference 10, particularly regarding the use of a damping function to minimise artefacts in the spectrum.

Our calculations are not aimed at reproducing the experimental spectrum to a high degree of accuracy; the relatively low-level electronic structure method used to generate the PES precludes such efforts. We note that the original MCTDH studies of pyrazine[15–17] used a PES which was finely tuned to reproduce the experimental spectrum. Instead, our aim here is to show that a reasonably-accurate absorption spectrum can be calculated (in a reasonable wall-time) using *ab initio* calculations and our SVD/additive-kernel method. The total wall time for our first calculation was 101 hours (including all electronic structure calculations) using a single processor on a desktop machine, whilst the second propagation took just $6\frac{1}{2}$ minutes. Around 88.6% of the first MCTDH calculation time was taken up in evaluating the KRR variance (Eq. (16)); clearly there is a need for improvement here if possible.

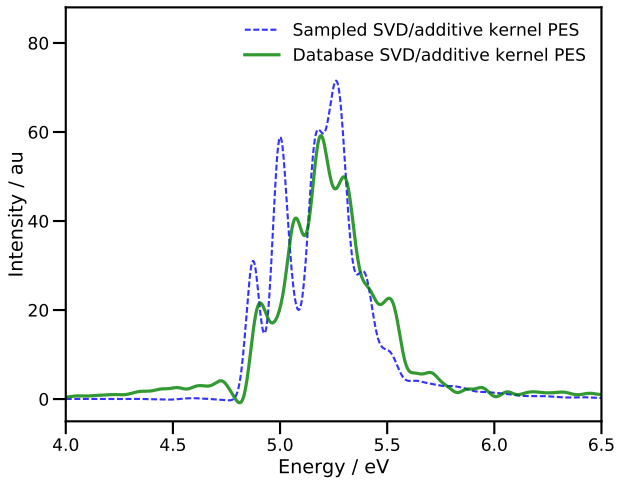

FIG. 4. Absorption spectra of pyrazine after vertical excitation to the S_2 state. Energies are relative to the S_0 minimum. The blue dashed line is the spectrum obtained in a DD-MCTDH calculation using an additive kernel and SVD fitting of the PES, with energies being added to the database on-the-fly. The green solid line is the spectrum from a DD-MCTDH calculation using the full database of energies generated during the first propagation.

The calculated spectra highlight the need to run an initial calculation (or even perhaps more than one in order to build up a converged PES) which samples configuration space before running a final calculation with a static PES database. The spectrum from the first propagation has some unphysical, negative regions (around 4.8 eV) along with tails at high and low energies, suggesting a low-accuracy representation of the PES. Furthermore, the addition of new energies to the PES database every femtosecond means that the PESs over which the wavepacket is moving are time-dependent, such that energy is not strictly conserved, leading to inaccuracies in the spectrum.

The conclusions from Fig. 4 are backed up by Fig. 5, which shows the population of the diabatic state corresponding to the S_2 adiabatic state at the Franck-Condon point. These results show significant population transfer between diabatic states, showing that our method can reproduce the diabatic couplings between states which are needed to model non-adiabatic effects such as the dynamics through conical intersections. Our calculations also show the immediate depopulation of the second excited state which was seen in the earlier MCTDH studies[15] indicating that, although the electronic structure method

used here is not highly-accurate, we are able to reproduce a key feature of the dynamics of pyrazine. The qualitative similarities in population transfer follow through the duration of the dynamics; the population reaches a minimum after 30 fs (after about 45 fs in the earlier work[15]) before a recurrence peaking after 60 fs (80 fs previously). These similarities indicate that, given further improvements in the efficiency in our algorithm, and using a more accurate electronic structure method, the DD-MCTDH method proposed here is capable of producing accurate results in a reasonable time.

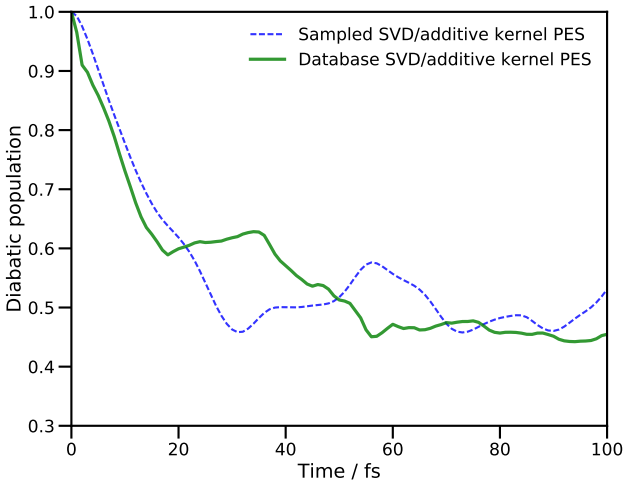

FIG. 5. Population of the second excited diabatic electronic state of pyrazine after vertical excitation. The blue dashed line is the result obtained by a DD-MCTDH calculation using an additive kernel and subsequent SVD fitting of the potential with energies being added to the database on-the-fly. The solid green line is the result of a DD-MCTDH calculation using the database of energies generated by the first propagation.

V. CONCLUSIONS

We have presented two improvements to our previously published DD method, both of which significantly improve the computational efficiency of the method: the use of an additive kernel in KRR greatly reduces the number of electronic structure calculations required to represent the PES, while the SVD fitting procedure can quickly and accurately reduce the number of terms in the resulting potential energy operator. Our simulations of proton transfer in salicylaldimine and non-adiabatic dynamics in pyrazine show that our method can be used to efficiently model quantum dynamics in many-dimensional systems.

That said, further developments are clearly warranted. As mentioned in the discussion of the pyrazine results, calculating KRR variance is a relatively expensive operation and should be addressed. Further work on the choice of the kernel is needed to determine whether an alternative would allow further reductions in the required number of electronic structure calculations. It is also possible to look into optimisation of the parameters in the kernel, particularly the coefficient of the exponent (which determines the width of the Gaussian functions); an optimal choice here will again reduce the number of electronic structure calculations needed. The use of gradient information to improve the representation of the PES is another possible avenue of investigation. We are also aware that the structure of a PES is not necessarily limited to interactions between, at most, two DOFs, and interactions between three or more DOFs may be required in an accurate expansion of the potential. Extension to include higher-order terms is ongoing, focusing on tensor decomposition to generalise the SVD fitting procedure. Finally, we note that the *a priori* choice of DOFs to include in the dynamics is an open question which we are aiming to address, particularly when dealing with larger molecules.

However, with the developments proposed here, an on-the-fly implementation of MCTDH, without the arduous task of pre-fitting a sum-of-products potential energy operator, is now a reality.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Leverhulme Trust for funding (RPG- 2016-055) and the Scientific Computing Research Technology Platform at the University of Warwick for providing computational resources. Data from Figs. 1-5 may be accessed at http://wrap.warwick.ac.uk/id/eprint/98400.

[1] N. Boekelheide, R. Salomón-Ferrer, and T. F. Miller III, Proc. Nat. Acad. Sci. USA **108**, 16159 (2011).
[2] L. Masgrau, A. Roujeinikova, L. O. Johannissen, P. Hothi, J. Basran, K. E. Ranaghan, A. J. Mulholland, M. J. Sutcliffe, N. S. Scrutton, and D. Leys, Science **312**, 237 (2006).
[3] C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández, and B. Kohler, Annu. Rev. Phys. Chem. **60**, 217 (2009).
[4] V. G. Stavros and J. R. Verlet, Annu. Rev. Phys. Chem. **67**, 211 (2016).
[5] L. A. Baker, M. D. Horbury, S. E. Greenough, F. Allais, P. S. Walsh, S. Habershon, and V. G. Stavros, J. Phys. Chem. Lett. **7**, 56 (2016).
[6] L. A. Baker, S. E. Greenough, and V. G. Stavros, J. Phys. Chem. Lett. **7**, 4655 (2016).
[7] A. H. Zewail, Science **242**, 1645 (1988).
[8] A. Douhal, S. K. Kim, and A. H. Zewail, Nature **378**, 260 (1995).
[9] A. H. Zewail, J. Phys. Chem. A **104**, 5660 (2000).
[10] M. H. Beck, A. Jäckle, G. A. Worth, and H. D. Meyer, Phys. Rep. **324**, 1 (2000).
[11] M. Schröder, F. Gatti, and H.-D. Meyer, J. Chem. Phys. **134**, 234307/1 (2011).
[12] M. Schröder and H.-D. Meyer, J. Chem. Phys. **141**, 034116/1 (2014).
[13] M. Coutino-Neto, A. Viel, and U. Manthe, J. Chem. Phys. **121**, 9207 (2004).
[14] K. Sadri, D. Lauvergnat, F. Gatti, and H.-D. Meyer, J. Chem. Phys. **141**, 114101/1 (2014).

[15]G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **105**, 4412 (1996).
[16]G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **109**, 3518 (1998).
[17]A. Raab, G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **110**, 936 (1999).
[18]S. Neville and G. Worth, J. Chem. Phys. **140**, 034317/1 (2014).
[19]T. J. Penfold and G. A. Worth, J. Chem. Phys. **131**, 064303/1 (2009).
[20]M. D. H. Köppel, I. Baldea, H.-D. Meyer, and P. Szalay, J. Chem. Phys. **117**, 2657 (2002).
[21]H. Wang and M. Thoss, J. Chem. Phys. **119**, 1289 (2003).
[22]H. Wang, J. Phys. Chem. A **119**, 7951 (2015).
[23]O. Vendrell and H.-D. Meyer, J. Chem. Phys. **134**, 044135/1 (2011).
[24]T. Hammer and U. Manthe, J. Chem. Phys. **134**, 224305/1 (2011).
[25]K. Giri, E. Chapman, C. S. Sanz, and G. Worth, J. Chem. Phys. **135**, 044311/1 (2011).
[26]H. Köppel, W. Domcke, and L. S. Cederbaum, Adv. Chem. Phys. **57**, 59 (1984).
[27]L. S. Cederbaum, H. Köppel, and W. Domcke, Int. J. Quant. Chem. **15**, 251 (1981).
[28]H.-D. Meyer, F. Gatti, and G. A. Worth, eds., *Multidimensional quantum dynamics: MCTDH theory and applications* (Wiley, Weinheim, Germany, 2009).
[29]M. Persico and G. Granucci, Theo. Chem. Acc. **133**, 1526/1 (2014).
[30]G. A. Worth, M. A. Robb, and B. Lasorne, Mol. Phys. **106**, 2077 (2008).
[31]J. C. Tully and R. K. Preston, J. Chem. Phys. **55**, 562 (1971).
[32]J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).
[33]M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, J. Phys. Chem. Lett. **3**, 3090 (2012).
[34]M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, J. Chem. Theory Comput. **7**, 1253 (2011).
[35]J. C. Tully, Farad. Discuss. **110**, 407 (1998).
[36]D. F. Coker, in *Computer Simulation in Chemical Physics*, edited by M. P. Allen and D. J. Tildesley (Kluwer Academic, Dordrecht, 1993) pp. 315–377.
[37]M. S. Topaler, M. D. Hack, T. C. Allison, Y.-P. Liu, S. L. Mielke, D. W. Schwenke, and D. G. Truhlar, J. Chem. Phys. **106**, 8699 (1997).
[38]B. R. Smith, M. J. Bearpark, M. A. Robb, F. Bernardi, and M. Olivucci, Chem. Phys. Lett. **242**, 27 (1995).
[39]M. J. Bearpark, F. Bernardi, M. Olivucci, M. A. Robb, and B. R. Smith, J. Am. Chem. Soc. **118**, 5254 (1996).
[40]R. Mitrić, J. Petersen, and V. Bonacić-Koutecký, Phys. Rev. A **79**, 053416/1 (2009).
[41]P. Lisinetskaya and R. Mitrić, Phys. Rev. A **83**, 033408/1 (2011).
[42]M. D. Hack, A. Jasper, Y. L. Volobuev, D. W. Schwenke, and D. G. Truhlar, J. Phys. Chem. A **103**, 6309 (1999).
[43]M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).
[44]M. Ben-Nun and T. J. Martínez, Adv. Chem. Phys. **121**, 439 (2002).
[45]H. R. Hudock, B. G. Levine, A. L. Thompson, H. Satzger, D. Townsend, N. Gador, S. Ullrich, A. Stolow, and T. J. Martínez, J. Phys. Chem. A **111**, 8500 (2007).
[46]T. Mori, W. Glover, M. Schuurman, and T. Martínez, J. Phys. Chem. A **116**, 2808 (2012).
[47]G. Worth and I. Burghardt, Chem. Phys. Lett. **368**, 502 (2003).
[48]I. Burghardt, H.-D. Meyer, and L. S. Cederbaum, J. Chem. Phys. **111**, 2927 (1999).
[49]G. Richings, I. Polyak, K. Spinlove, G. Worth, I. Burghardt, and B. Lasorne, Int. Rev. Phys. Chem. **34**, 269 (2015).
[50]G. A. Worth, M. A. Robb, and I. Burghardt, Farad. Discuss. **127**, 307 (2004).
[51]B. Lasorne, M. J. Bearpark, M. A. Robb, and G. A. Worth, J. Phys. Chem. A **112**, 13017 (2008).
[52]B. Lasorne, M. A. Robb, and G. A. Worth, Phys. Chem. Chem. Phys. **9**, 3210 (2007).
[53]D. Mendive-Tapia, B. Lasorne, G. Worth, M. Bearpark, and M. Robb, Phys. Chem. Chem. Phys. **12**, 15725 (2010).
[54]C. Allan, B. Lasorne, G. Worth, and M. Robb, J. Phys. Chem. A **114**, 8713 (2010).
[55]D. Asturiol, B. Lasorne, G. Worth, M. Bearpark, and M. Robb, Phys. Chem. Chem. Phys. **12**, 4949 (2010).
[56]M. Araújo, B. Lasorne, A. Magalhaes, M. Bearpark, and M. Robb, J. Phys. Chem. A **114**, 12016 (2010).
[57]I. Polyak, C. Allan, and G. Worth, J. Chem. Phys. **143**, 084121/1 (2015).
[58]M. Vacher, M. J. Bearpark, and M. A. Robb, Theor. Chem. Acc. **135**, 187/1 (2016).
[59]G. Richings and S. Habershon, Chem. Phys. Lett. **683**, 228 (2017).
[60]G. Richings and S. Habershon, J. Chem. Theory Comput. **13**, 4012 (2017).
[61]C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, Cambridge, Massachusetts, 2006).
[62]C. Williams, in *Handbook of Brain Theory and Neural Networks*, edited by M. Arbib (The MIT Press, Cambridge, Massachusetts, 2002) pp. 466–470.
[63]A. Bartók and G. Csányi, Int. J. Quantum. Chem. **115**, 1051 (2015).
[64]J. P. Alborzpour, D. P. Tew, and S. Habershon, J. Chem. Phys. **145**, 174112/1 (2016).
[65]L. Mones, N. Bernstein, and G. Csányi, J. Chem. Theory Comput. **12**, 5100 (2016).
[66]J. Quinonero-Candela, C. E. Ramussen, and C. K. I. Williams, in *Large-Scale Kernel Machines*, edited by L. Bottou, O. Chapelle, D. DeCoste, and J. Weston (The MIT Press, Cambridge, Massachusetts, 2007) pp. 203–224.
[67]D. Duvenaud, H. Nickisch, and C. E. Rasmussen, in *Neural Information Processing Systems Conference* (2011).
[68]K. Chalupka, C. K. Williams, and I. Murray, J. Mach. Learn Res. **14**, 333 (2013).
[69]P. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).
[70]J. Frenkel, *Wave Mechanics* (Clarendon Press, Oxford, 1934).
[71]G. Worth, K. Giri, G. Richings, I. Burghardt, M. Beck, A. Jäckle, and H.-D. Meyer, "The Quantics Package, Version 1.1," Tech. Rep. (University of Birmingham, Birmingham, UK, 2016).
[72]G. Richings and G. Worth, J. Phys. Chem. A **119**, 12457 (2015).
[73]G. Richings and G. Worth, Chem. Phys. Lett. **683**, 606 (2017).
[74]B. Esry and H. Sadeghpour, Phys. Rev. A **68**, 042706/1 (2003).
[75]M. Baer, Chem. Phys. Lett **35**, 112 (1975).
[76]H. Rabitz and O. Aliş, J. Math. Chem. **25**, 197 (1999).
[77]O. Aliş and H. Rabitz, J. Math. Chem. **29**, 127 (2001).
[78]T.-S. Ho and H. Rabitz, J. Chem. Phys. **119**, 6433 (2003).
[79]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, "Molpro, version 2015.1, a package of ab initio programs," (2015).
[80]H.-J. Werner and P. J. Knowles, J. Chem. Phys. **82**, 5053 (1985).

MCTDH on-the-fly: Efficient grid-based quantum dynamics without pre-computed potential energy surfaces

Gareth W. Richings[1, a)] and Scott Habershon[1, b)]
Department of Chemistry and Centre for Scientific Computing, University of Warwick, Coventry, CV4 7AL, UK

We present significant algorithmic improvements to a recently-proposed direct quantum dynamics method, based upon combining well established grid-based quantum dynamics approaches and expansions of the potential energy operator in terms of a weighted sum of Gaussian functions. Specifically, using a sum of low-dimensional Gaussian functions to represent the potential energy surface (PES), combined with a secondary fitting of the PES using singular value decomposition, we show how standard grid-based quantum dynamics methods can be dramatically accelerated without loss of accuracy. This is demonstrated by on-the-fly simulations (using both standard grid-based methods and MCTDH) of both proton transfer on the electronic ground state of salicylaldimine and the non-adiabatic dynamics of pyrazine.

I. INTRODUCTION

Many chemical processes require a quantum-mechanical description to provide the full picture of the molecular dynamics; examples include proton tunnelling in enzyme catalysis[1,2], photochemistry of DNA molecules[3,4] or plant sunscreens[5,6], and other aspects of femtochemistry[7–9]. The most direct route to performing quantum dynamical simulations is to provide solutions to the nuclear, time-dependent Schrödinger equation (TDSE)

$$i\hbar\frac{\partial\Psi\left(\mathbf{q},t\right)}{\partial t}=\hat{H}\Psi\left(\mathbf{q},t\right), \tag{1}$$

where $\mathbf{q}$ is a vector representing the nuclear degrees-of-freedom (DOFs) of the system in question. Once the wavefunction, $\Psi(\mathbf{q},t)$ is known then one can extract any information required about the time-evolution of the system.

Solution of the TDSE using so-called grid-based methods is now quite standard; here, the wavefunction is expanded and evolved on a grid of localised basis functions, regularly distributed through configuration space[10]. The multi-configuration time-dependent Hartree (MCTDH) method,[10–14] a natural extension of grid-based methods which simultaneously employs time-dependent basis functions, has become the *de facto* gold standard in the area of wavefunction dynamics. MCTDH is capable of treating a few tens of DOFs and converges to the numerically exact solution of the TDSE (given a suitable basis of sufficient size). MCTDH has been used to study, for example: the photochemistry of pyrazine[15–17], pyrrole[18] and benzene[19,20]. The upper limit on the system size treatable by MCTDH has been extended by the introduction of the multi-layer MCTDH method[21–24], bringing systems of several hundred DOFs into the purview of fully quantum simulations.

However, there remain significant barriers to using grid-based quantum dynamics methods which mean that their more general uptake by wider computational and experimental communities remains low. The key challenge of using grid-based quantum dynamics is the requirement that the system Hamiltonian must be fully defined before performing wavepacket propagation. First, only a subset of the DOFs of the molecular system are studied to make the dynamics calculation tractable, so one is faced with a choice as to which DOFs might be important; this is an important challenge when encountering a previously-unstudied system. Second, because quantum mechanics is non-local, the global form of the PES must be known prior to performing the dynamics. In simulations using grid-based methods such as MCTDH, the global PES is usually determined by fitting pre-defined functions to a set of electronic energies calculated at various molecular geometries, for example by using the vibronic-coupling Hamiltonian (VCHAM) method which is based on PES expansion in low-order polynomials[18,25–27]. This fitting of the PES is an extremely arduous task which must be performed before one can even begin to perform MCTDH simulations; this is the central barrier to making grid-based methods like MCTDH more widely applicable. As an aside, it is worth noting that a similar difficulty can arise for the kinetic energy operator, although this can often be addressed by an appropriate choice of (*e.g.* rectilinear) coordinates before commencing PES fitting.[28]

To make accurate, grid-based quantum dynamics methods (*e.g.* MCTDH) more accessible and more widely applicable, our recent efforts have focused on coupling aspects of direct-dynamics (DD) simulations with standard-grid-based approaches. In DD methods, the PES is calculated "on-the-fly" during propagation of the dynamical wavefunction;[29,30] examples include trajectory surface hopping (TSH[31–42]), *ab initio* multiple spawning (AIMS[43–46]), the DD-variational multi-configuration Gaussian (DD-vMCG) method[47–58], and related Gaussian wavepacket-based approaches. These approaches all have the advantage of only requiring PES evaluations where and when they are needed during

a) Electronic mail: G.Richings@warwick.ac.uk
b) Electronic mail: S.Habershon@warwick.ac.uk

wavefunction propagation. However, these DD methods also have features which can limit their domain of application; for example, TSH and AIMS rely on classical equations-of-motion (EOMs) to propagate trajectories or basis functions, such that tunnelling and zero-point energy conservation cannot be treated explicitly, while the (fully quantum-mechanical) DD-vMCG method can suffer from numerical instability and linear dependence in the solution of the EOMs due to the non-orthogonality of the Gaussian basis functions.[49]

In contrast, our recently-proposed methodology combines the stability of grid-based quantum dynamics, particularly MCTDH, with the convenience of DD methods[59,60]. Our method relies on the ideas of Gaussian process regression (GPR) and kernel ridge regression (KRR)[61–68] to express the PES as a weighted sum of Gaussian functions; we refer to our overall approach hereafter as the DD-grid-based (DD-GB) method (with specific names, defined later, used when employing a particular quantum propagation method). Most importantly, the Gaussian-kernel-based representation of the PES used in our simulation strategy is capable of constructing a global PES using just a few local PES evaluations, and is therefore employable as a DD strategy; in addition, the results Gaussian-based representation of the PES is already in the sum-of-products form which is required for efficient MCTDH propagation. Furthermore, as discussed below, this DD-GB strategy is also compatible with on-the-fly strategies for dealing with non-adiabatic transitions, and so may be used to simulate photochemical dynamics as well as ground-state dynamics.

Our DD-GB method has already been shown to give accurate dynamical results when compared to benchmark studies using both fitted and *ab initio* on-the-fly PESs; however, our initial investigations also noted that the existing approach has significant computational demands associated with PES construction and wavefunction propagation. The aim of this work is to demonstrate two new algorithmic improvements to our DD-GB method which dramatically decrease computational effort, and hence greatly increase the potential utility of our approach. First, in Section III A, we show how additive kernels can be exploited in our KRR PES representation to reduce the number and dimensionality of integrals required for wavefunction propagation on a grid. Second, in Section IV, we show how singular value decomposition (SVD) can be used to generate a compact representation of KRR PESs generated on-the-fly during wavefunction propagation. Together, these two extensions are demonstrated to yield versions of GB and MCTDH propagations which are both accurate and essentially DD in nature; this is highlighted in DD-GB simulations of proton transfer in salicylaldimine and non-adiabatic dynamics of pyrazine.

II. METHODOLOGY

A. Grid-Based Quantum Dynamics: The Standard Method

The standard grid-based method for quantum dynamics is well established,[10] and we give here only the details relevant to the simulations presented below. For a molecular system with f DOFs, the wavefunction is expanded in terms of products of *time-independent* basis functions, each of which has an associated, time-dependent, complex coefficient, $C_{j_1,\cdots,j_f}$. For a nuclear wavepacket moving on electronic state, s, we then have

$$\begin{aligned}\Psi^{(s)}(q_1,\cdots,q_f,t) &= \sum_{j_1}^{N_1}\cdots\sum_{j_f}^{N_f} C^{(s)}_{j_1,\cdots,j_f}(t)\prod_{\kappa=1}^{f}\chi^{(\kappa)}_{j_\kappa}(q_\kappa)\\ &= \sum_{J} C^{(s)}_{J}(t)\,X_J(\mathbf{q})\end{aligned} \tag{2}$$

Note that we have taken the opportunity to introduce a compound index, $J{=}j_1,\cdots,j_f$. The total wavefunction for a system with N_s orthonormal electronic states is defined as

$$|\Psi\rangle = \sum_{s=1}^{N_s}|\Psi^{(s)}\rangle|s\rangle \tag{3}$$

Writing the total Hamiltonian as

$$\hat{H} = \sum_{su}^{N_s}|s\rangle\hat{H}^{(su)}\langle u|, \tag{4}$$

and employing the Dirac-Frenkel variational principle (DFVP),[69,70] we get a set of coupled EOMs for the expansion coefficients:

$$i\hbar\dot{C}^{(s)}_J = \sum_{u}^{N_s}\sum_{L}\langle X_J|\hat{H}^{(su)}|X_L\rangle C^{(u)}_L \tag{5}$$

Integration of these EOMs allows one to follow the propagation of the wavefunction through time. Given an appropriate basis of sufficient size, the propagation of these EOMs provides a numerically exact solution of the TDSE for a given Hamiltonian, and, as such, this method is referred to as the standard method (SM)[10].

B. Grid-Based Quantum Dynamics: MCTDH

Because of the exponential scaling with system size, the SM is restricted to treating about 5 DOFs in practical calculations. To enable simulations of larger molecular systems, MCTDH was introduced. In MCTDH, as in the SM, the wavefunction is expanded in a sum of products of

basis functions, each product having a complex expansion coefficient,

$$\begin{aligned}\Psi^{(s)}(Q_1,\cdots,Q_f,t)&\\ =\sum_{j_1}^{n_1}\cdots\sum_{j_m}^{n_m}A^{(s)}_{j_1,\cdots,j_m}(t)\prod_{\kappa=1}^{m}\varphi^{(s,\kappa)}_{j_\kappa}(Q_\kappa,t)&\\ =\sum_J A^{(s)}_J(t)\,\Phi^{(s)}_J(\mathbf{Q},t)\,.&\end{aligned}\tag{6}$$

The key difference between the SM and MCTDH *ansätze* is that the basis functions in the MCTDH method are *time-dependent*. Combining the MCTDH *ansatz* and DFVP yields EOMs for the coefficients and the time-dependent basis functions (referred to as single-particle functions (SPFs)):

$$i\hbar\dot{A}^{(s)}_J=\sum_u^{N_s}\sum_L\langle\Phi^{(s)}_J|\hat{H}^{(su)}|\Phi^{(u)}_L\rangle A^{(u)}_L\tag{7a}$$

$$i\hbar\dot{\boldsymbol{\varphi}}^{(s,\kappa)}=\left(1-P^{(s,\kappa)}\right)\left(\boldsymbol{\rho}^{(s,\kappa)}\right)^{-1}\sum_u^{N_s}\langle\hat{H}^{(su)}\rangle^{(\kappa)}\boldsymbol{\varphi}^{(u,\kappa)}\tag{7b}$$

The SPFs $\boldsymbol{\varphi}^{(s,\kappa)}$ are functions of a small subset (usually 1-4) of the system DOFs, $Q_\kappa=(q_{\kappa_1},\cdots,q_{\kappa_p})$, and are themselves expansions in terms of a time-independent basis (as in the SM wavefunction in Eq. (2)):

$$\varphi^{(s,\kappa)}_{j_\kappa}(Q_\kappa,t)=\sum_{i_\kappa}^{N_\kappa}c^{(s,\kappa,j_\kappa)}_{i_\kappa}(t)\,X^{(\kappa)}_{i_\kappa}(Q_\kappa)\,.\tag{8}$$

In the SPF EOMs (Eq. (7b)), $\hat{P}^{(s,\kappa)}$ is a projection operator onto the SPF space along mode κ, and $\left(\boldsymbol{\rho}^{(s,\kappa)}\right)^{-1}$ is the inverse of the density matrix associated with κ. By constructing a Hartree product of SPFs in all modes apart from κ, $\Phi^{(s)}_{J^\kappa}$, we can define a set of single-hole functions, $\Psi^{(s,\kappa)}_l=\sum_{J^\kappa}A^{(s)}_{J^\kappa_l}\Phi^{(s)}_{J^\kappa}$, from which a mean-field matrix, with elements $\langle\hat{H}^{(st)}\rangle^{(\kappa)}_{jl}=\langle\Psi^{(s,\kappa)}_j|\hat{H}^{(st)}|\Psi^{(t,\kappa)}_l\rangle$, is defined.

By evolving the SPFs variationally, it is possible to keep their number to a minimum; it is this reduction in the size of the basis in Eq. (6) which permits the study of larger systems than could be treated by the SM (MCTDH also scales exponentially, but with a smaller base than the SM[10]).

C. PESs for Grid-Based Wavefunction Propagation

To solve the EOMs in Eqs. (5) and (7) as efficiently as possible, the Hamiltonian must be in a sum-of-products form (*i.e.* all terms are products of functions of single DOFs), so that the multi-dimensional integrals can be reduced to sums-of-products of one-dimensional integrals. This feature of the Hamiltonian is not usually a problem for the kinetic energy part if a sensible choice of coordinate system is made, but this requirement can be difficult to ensure for the potential energy part of the Hamiltonian.

In recent work[59,60] we have presented a grid-based DD method using Eqs. (5) and (7) where the PES is constructed on-the-fly by using KRR fit to a set of PES values calculated at appropriately chosen molecular geometries, $\{\mathbf{q}^k\}$, at which the Gaussian kernel functions of the PES are also centered. Defining a one-dimensional kernel function along DOF, κ, as

$$k(q_\kappa,q^k_\kappa)=\exp(-\alpha(q_\kappa-q^k_\kappa)^2)\tag{9}$$

the potential energy operator is then represented as

$$\begin{aligned}V^{(su)}(\mathbf{q})&\approx\sum_{k=1}^{M}w^{(su)}_k\prod_{\kappa=1}^{f}k(q_\kappa,q^k_\kappa)\\ &=\sum_{k=1}^{M}w^{(su)}_k k(\mathbf{q},\mathbf{q}^k)\end{aligned}\tag{10}$$

The width parameter, α, can in principle be optimized by log-likelihood maximization, as employed in GPR, although we have found to date that using an appropriate fixed value of α, chosen at the outset, is sufficient for our purposes. The weights of the expansion, $\{w^{(su)}_k\}$, are determined by solution of the linear equation

$$\mathbf{K}\mathbf{w}=\mathbf{b},\tag{11}$$

where the vector, $\mathbf{w}$, contains the weights and the elements of the covariance matrix, $\mathbf{K}$, are[62]

$$K_{ij}=k(\mathbf{q}^i,\mathbf{q}^j)+\gamma^2\delta_{ij},\tag{12}$$

with γ, being a small, positive regularisation parameter. The vector, $\mathbf{b}$, contains the PES values calculated (*e.g.* using *ab initio* electronic structure calculations) at sampled geometries,

$$b_i=V^{(su)}\left(\mathbf{q}^i\right).\tag{13}$$

Returning to Eq. (10), we find that, if we use rectilinear coordinates (such as Cartesian coordinates or normal mode coordinates) in the dynamics calculation, the representation of the potential energy operator has the sum-of-products form necessary for efficient integral evaluation,

$$\langle X_J|\hat{V}^{(su)}|X_L\rangle=\sum_{k=1}^{M}w^{(su)}_k\prod_{\kappa=1}^{f}\int\mathrm{d}q_\kappa\,\chi^*_{j_\kappa}k(q_\kappa,q^k_\kappa)\chi_{l_\kappa}.\tag{14}$$

The individual integrals can be further simplified by an appropriate choice of time-independent basis; the implementation of the SM and MCTDH in the Quantics code[71] uses discrete variable representation (DVR) basis

functions[10], which are centered at points in configuration space, $\mathbf{q}_\alpha$ yielding a grid along each DOF, such that

$$\langle\chi_\alpha|\hat{V}^{(su)}|\chi_\beta\rangle \approx V\left(\mathbf{q}_\alpha\right)\delta_{\alpha\beta}. \tag{15}$$

This property means that the integrals in Eq. (14) can be evaluated simply by using the value of each term in the potential energy function at the location of each DVR gridpoint.

The KRR-based approach outlined above requires one to select the molecular geometries used in the PES fit. In DD-GB, this is implemented by randomly sampling a set number of geometries around the centre of the wavepacket at pre-determined intervals, and then using those geometries to calculate PES reference values. However, to reduce computational effort through calculating excessive numbers of potential energies, it is possible to avoid placing new reference points in regions of configuration space where the PES representation is already sufficiently accurate, we first compute the variance function[62]

$$\sigma^2\left(\mathbf{q}\right) = k(\mathbf{q},\mathbf{q}) + \gamma^2 - \mathbf{k}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{k}. \tag{16}$$

at the randomly-sampled geometries, $\mathbf{q}$. Here, γ and $\mathbf{K}$ are defined as in Eq. (12) whilst the vector, $\mathbf{k}$, contains the covariances of the new point with all of the pre-existing points (Eq. (9)). To avoid unnecessary PES evaluations in our scheme, we define a numerical parameter which, if exceeded by the variance at $\mathbf{q}$, signals that a new energy is required to be calculated; otherwise, the selected geometry is rejected and another sampled.

The key feature of this KRR-based methodology is that it brings GB/MCTDH simulations closer to the DD approach. One does not need to pre-fit or pre-compute a global PES; instead, the potential energy operator is learnt on-the-fly using PES evaluations (usually *ab initio* calculations) at configurations sampled from the evolving wavepacket. As a final point, we note that, after an initial wavepacket propagation and PES construction simulation has been performed, one can restart the simulation from the initial wavepacket using the full KRR-PES to obtain a final picture of the quantum dynamics.

1. Non-Adiabatic Simulations

To study multiple electronic states, and the non-adiabatic transitions between them, using grid-based methods, it is desirable to use a quasi-diabatic PES where the discontinuities in the gradient of the adiabatic surfaces and the non-adiabatic couplings are transformed away, yielding smoothly crossing surfaces. To achieve such a transformation on-the-fly, we use a modified version of a scheme proposed by one of us[72,73] in the context of DD-vMCG simulations. This approach is based on propagation of the diabatization matrix,[74] $\mathbf{A}$, using line integrals of the non-adiabatic coupling terms between the adiabatic states, ψ_i and ψ_j (with respective energies V_{ii}^A and V_{jj}^A), given by

$$\mathbf{F}_{ij} = \frac{\langle\psi_i|\nabla\hat{H}|\psi_j\rangle}{V_{jj}-V_{ii}}. \tag{17}$$

The approximate relationship[75]

$$\nabla\mathbf{A} \approx -\underline{\mathbf{F}}\mathbf{A}, \tag{18}$$

holds for an incomplete set of adiabatic states (it is an equality for a complete set) where the underlining of $\underline{\mathbf{F}}$ indicates that it is a matrix of vectors. As we are interested in non-radiative transfers between a small number of (usually two) energetically-close states, we take the pragmatic view of using Eq. (18) as an equality in the method here. Starting at a geometry, $\mathbf{q}$, where $\mathbf{A}$ is known, Eq. (18) is integrated between $\mathbf{q}$ and some new point, $\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}$, to give $\mathbf{A}$ at the new point:

$$\mathbf{A}\left(\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}\right) = \exp\left(-\int_{\mathbf{q}}^{\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}}\underline{\mathbf{F}}\cdot\mathrm{d}\mathbf{q}\right)\mathbf{A}\left(\mathbf{q}\right). \tag{19}$$

The matrix, $\mathbf{A}$, then allows transformation of the adiabatic energy matrix, $\mathbf{V}^{\mathbf{A}}$, to the quasi-diabatic representation

$$\mathbf{V}^{\mathrm{D}} = \mathbf{A}^{\mathrm{T}}\mathbf{V}^{\mathrm{A}}\mathbf{A}. \tag{20}$$

In practice the choice of $\mathbf{A}$=$\mathbf{I}$ is made at the center of the initial wavepacket and the transformation matrix propagated away from this point towards newly-sampled geometries as the evolution of the wavepacket and PES proceeds.

2. Problems with Computational Efficiency

Previous work[59,60] has already demonstrated that the scheme outlined above can reproduce the quantum dynamics results obtained by using pre-fitted PESs or DD-vMCG. However, we have also noted[60] the lack of efficiency of our original scheme, especially when coupled to MCTDH propagation of the wavefunction; even for a system comprising six DOFs and one electronic state (a modest system by normal MCTDH standards), we found that a propagation of just 100 fs could take several weeks to perform. Such time requirements clearly make the method, as described above, of limited applicability.

The inefficiency of our original DD-MCTDH scheme is caused by the large number of reference points needed to expand the PES (*i.e.* the large value of M in Eq. (14)). Assuming an MCTDH wavefunction constructed from one-dimensional SPFs (the arguments presented also follow for wavefunctions constructed using multi-dimensional SPFs), the PES contribution to the Hamiltonian integral in Eq. (7a), when using the KRR-based approach outlined above is then

$$\begin{aligned}&\langle\Phi_J^{(s)}|\hat{V}^{(su)}|\Phi_L^{(u)}\rangle \\ &\qquad= \sum_{k=1}^{M} w_k^{(su)}\prod_{\kappa=1}^{f}\int\mathrm{d}q_\kappa\varphi_{j_\kappa}^{(s,\kappa)*}k(q_\kappa,q_\kappa^k)\varphi_{l_\kappa}^{(u,\kappa)}\end{aligned} \tag{21}$$

Inserting Eq. (8) we get

$$\langle \Phi_J^{(s)} | \hat{V}^{(su)} | \Phi_L^{(u)} \rangle = \sum_{k=1}^{M} w_k^{(su)} \prod_{\kappa=1}^{f} \sum_{m_\kappa, n_\kappa}^{N_\kappa} c_{m_\kappa}^{(s,\kappa,j_\kappa)*} c_{n_\kappa}^{(u,\kappa,l_\kappa)} \times \int \mathrm{d}q_\kappa \chi_{m_\kappa}^{*} k(q_\kappa, q_\kappa^k) \chi_{n_\kappa} \tag{22}$$

Because the SPFs are time-dependent the transformations of the DVR integrals must be performed at *every* step of wavefunction propagation, an extremely time-consuming process. In an ideal world, the transformation to the SPF basis would be performed on the full DVR potential integral, in other words sum up the DVR integrals for all potential terms, then transform. To do so would require the two summations in Eq. (22) to swap place, or as the summation over m_κ, n_κ depends on the product index, at least for the GPR summation to be moved to the inner position. However, this is clearly not a valid change, so we are stuck with this time-consuming form. The problem is even worse when considering the mean-field matrices in Eq. (7b) because a large number of such matrices must be constructed in a similar way to Eq. (22).

Frustratingly, calculations using the SM do not suffer from this efficiency bottleneck because there is no need to perform a transformation to a time-dependent basis (see Eq. (5)). As described in the previous section, the potential energy operator is only updated occasionally during wavefunction propagation; once this has occurred the Hamiltonian integral in Eq. (5) can be evaluated, stored and used until the next update. This reflects the simplicity of the SM; it is only necessary to know the total value of the potential at each point on this product grid, not the separate contribution from each term, as required by MCTDH.

In summary, it is apparent that the best way to speed up DD-MCTDH calculations using the KRR expansion of the PES is to reduce the number of terms needed (obviously, without losing accuracy). Such a reduction will have the added benefit that fewer *ab initio* electronic structure calculations need to be performed. In the next section, we present an approach for achieving the reduction in the size of the KRR database.

III. METHODOLOGY IMPROVEMENT I: ADDITIVE KERNELS

A. Theory and Implementation

As outlined above, initial implementation of the DD-GB method used f-dimensional Gaussian functions as the kernels to fit to the PES (Eq. (10) and (12)); we refer to this kernel hereafter as the *full* kernel. The large number of such functions required to represent a PES is due to the Gaussian function's locality; for a given value of α, the full-width at half-minimum (FWHM) of the Gaussian along any DOF is $2\sqrt{\alpha \ln 2}$. It follows that a reasonable measure of the region of influence of each Gaussian is given by the volume of the hypersphere of that radius,

$$\text{Volume} = \frac{\pi^{f/2}}{\Gamma(f/2+1)} 2^f (\alpha \ln 2)^{f/2}. \tag{23}$$

Given some representative length, l, of the DVR grids, the volume of configuration space in which the wavepacket propagation proceeds is thus l^f, meaning that each Gaussian kernel "occupies" a fraction of configurational space given by

$$\mathrm{P} = \frac{1}{\Gamma(f/2+1)} \left(\frac{2\,\pi^{1/2} (\alpha \ln 2)^{1/2}}{l} \right)^f \tag{24}$$

Because the Gaussian widths are less than the grid length along each DOF, it is clear that the proportion of configuration space influenced by each Gaussian kernel decreases exponentially with increasing number of DOFs. To reduce the number of Gaussian functions required to span configuration space, it is thus necessary to use a less localised kernel. Fortunately, the choice of an f-dimensional Gaussian kernel is just one of many possible within GPR and KRR[61], with the main restriction on the choice of kernel being that it must ensure the covariance matrix, $\mathbf{K}$, is symmetric and positive semi-definite. To this end, it has been shown that a kernel constructed from a sum of lower dimensional Gaussian functions is a valid covariance function, as in

$$k^{\mathrm{Add}}(\mathbf{q}, \mathbf{q}^n) = \sum_{\kappa=1}^{f} k(q_\kappa, q_\kappa^n) + \sum_{\kappa,\lambda}^{f} k(q_\kappa, q_\kappa^n) k(q_\lambda, q_\lambda^n) + \cdots + k(\mathbf{q}, \mathbf{q}^n). \tag{25}$$

This type of expansion, termed an additive kernel, is reminiscent of that used within the high-dimensional model representation (HDMR)[76–78], a fit being built up from a set of increasingly higher-dimensional functions. Fitting a PES to a set of low-dimensional functions is also the approach used within the successful VCHAM approach, so such an approximation is valid when trying to fit multi-dimensional PESs, as well as fulfilling the requirement of being in the sum-of-products form for efficient MCTDH propagation. In practice, the additive kernel is simply used to replace the f-dimensional Gaussian functions in Eqs. (10) and (12); all other aspects of the KRR process remain as before.

How does such a kernel help in reducing the number of reference points required to construct a PES expansion? Returning to the volume of influence of the kernel function, consider one of the 1-dimensional terms, $\exp(-\alpha(q_\kappa - q_\kappa^n)^2)$. At some coordinate, q_κ^0, this term has the value $k(q_\kappa^0, q_\kappa^n)$, but as it is independent of the other $f-1$ coordinates it has the same value everywhere

within the configuration subspace defined as consisting of the points $(q_1, \cdots, q_\kappa^0, \cdots, q_f)$. The subspace has a volume of l^{f-1} therefore, based on the FWHM argument, the kernel has influence over a volume of $2\sqrt{\alpha \ln 2}\, l^{f-1}$; expressed as a proportion of the total configuration space, this volume is $2\sqrt{\alpha \ln 2}/l$. This volume is independent of the dimensionality of the problem at hand, so each reference point can describe a much larger proportion of the PES than if a single f-dimensional Gaussian was used. Similar arguments hold for all of the terms in Eq. (25) (except the final, f-dimensional term).

One can truncate the expansion in Eq. (25) at any order and still be left with a valid kernel; however, because two-body terms tend to dominate in expressions of the PES,[76] we propose here to use a two-body additive kernel (*i.e.* truncating the expansion in Eq. (25) after the second term) as an alternative to the f-dimensional Gaussian functions used previously. Extending our previous work,[59,60] we have implemented DD-SM and DD-MCTDH codes which employ a second-order additive kernel in a development version of the Quantics quantum dynamics package[71]. In the next section we present the results of calculations, using this implementation, which demonstrate the ability of the additive kernel to accurately represent the underlying PES with lower computational effort than the corresponding f-dimensional version employed in our previous work.

B. Salicylaldimine Proton Transfer

In this section, we present the results of wavefunction propagations performed on a 4-dimensional model of proton transfer in salicylaldimine using a development version of the Quantics quantum dynamics package[71]. Three different calculations were performed: (i) a reference SM calculation propagation using the VCHAM-fitted PES of Polyak *et al*[57], (ii) DD-SM on the same PES using a 4-dimensional kernel, and (iii) DD-SM on the same PES using a second-order additive kernel.

All calculations were performed in mass-frequency scaled normal mode coordinates; using the nomenclature of Polyak *et al* the in-plane modes ν_1 (proton transfer mode), ν_{13} (bending of O and N away from one another), ν_{32} (CO stretch and NH bend) and ν_{36} (OHN bending) were chosen as the system DOFs. A sine DVR of 101 gridpoints was used for the ν_1 mode whilst 21 member harmonic oscillator DVR bases were used for the other three modes. The initial wavepacket was a 4-dimensional Gaussian function centred at $\langle \nu_1 \rangle$=0.96 (to the enol side of the barrier), $\langle \nu_{13} \rangle = \langle \nu_{32} \rangle$=0 and $\langle \nu_{36} \rangle$=0.14, with widths of $\langle \mathrm{d}\nu_1 \rangle$=0.5706, $\langle \mathrm{d}\nu_{13} \rangle$=0.6902, $\langle \mathrm{d}\nu_{32} \rangle$=0.6707, $\langle \mathrm{d}\nu_{36} \rangle$=0.7704. Time-propagation was carried out for 100 fs; the DD-SM calculations sampled the PES every 1 fs, with 300 geometries being randomly selected within 3 times the width of the wavepacket from the centre in each DOF. PES values were calculated at the chosen geometries if the variance there was greater than 10^{-3}. A kernel width parameter, α, of 0.5 was used for all DD-SM calculations. Subsequently, DD-SM calculations were performed using the database of energies calculated previously, with the potential energy operator generated at the start of the calculation and with no further enlargement of the initial PES database. In all cases, the EOMs were integrated using the 15$^{\text{th}}$-order short iterative Lanczos (SIL) method with accuracy cutoff of 10^{-6}. A flux operator[10] was evaluated at q_1=0 during the propagation, allowing measurement of the flow of the wavefunction across the potential barrier for the proton transfer.

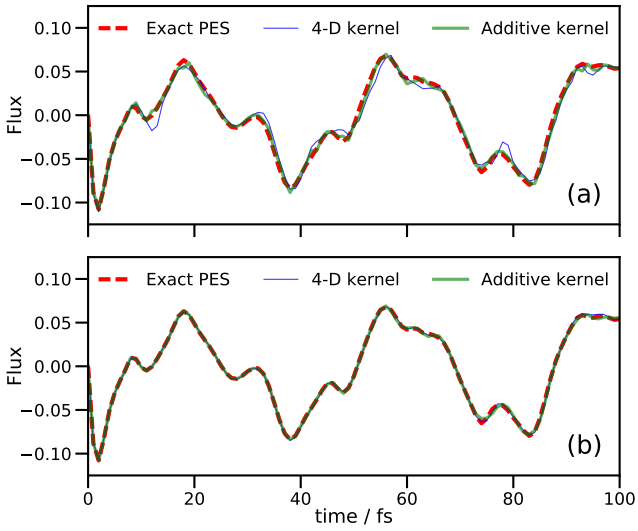

FIG. 1. Flux operator expectation value for a model of salicylaldimine with 4 DOFs. The dividing surface is placed at the barrier of the transition mode, ν_1. Both (a) and (b) show results of grid-based quantum dynamics using the standard grid method on the fitted PES (red dashed lines), a DD-SM simulation using a 4-dimensional KRR PES (thin blue line), and a DD-SM simulation using a second-order additive kernel (thick green line). In (a), DD-SM results are shown for the initial wavefunction propagation; in (b), DD-SM results are shown for a second simulation in which all reference points learned in (a) are used to construct the respective potential energy operators.

Figure 1(a) shows the flux operator expectation value along mode ν_1 for the full SM calculation and the two DD-SM calculations with different kernels; here, the energy database was built on-the-fly during the DD-SM calculations. All plots are qualitatively similar, with initial motion of the wavepacket across the barrier towards the keto tautomer, followed by the oscillation between the keto and enol sides of the barrier oscillates with a period of about 40 fs. The actual dynamics is as expected from previous work; the important point to note here is that the dynamics matches the reference SM results extremely well when using both the full kernel or the second-order additive kernel, with just small differences

appearing later on as inaccuracies due to the sampling of the potential creep in.

The key difference in performance between the kernels is in the calculation timings. Both DD-SM calculations employed a single core on a standard desktop machine; however, whilst the full-kernel calculation required 15.5 CPU hours, the additive kernel calculation took 2.25 CPU hours (an improvement by a factor of around 7). The origin of the difference in computational effort is the different number of points added to the database of electronic energies; the full kernel calculation generated 8,026 energies, but the additive kernel only required 741 geometries to be sampled to generate an accurate PES representation. This saving in the number of energy datapoints required in the two DD-GB methods accounts for nearly all of the difference in the computation time (in fact, the times required for wavefunction propagation were quite similar, 2,425 against 2,151 CPU seconds for the additive and full kernels, respectively). Furthermore, the calculation of the KRR variances and weights took nearly 10,000 CPU seconds with the full kernel, but only 36 CPU seconds with the additive kernel (this huge difference reflects the relative speeds of solving the linear equations in Eqs. (11) and (16) when using a matrix with 100 times fewer entries). Finally, the largest difference in timings was actually in calculating and storing the integrals in Eq. (15): 43,700 CPU seconds for the full kernel against 5,300 for the additive kernel. Both calculations used the same size DVR grid, so the difference in these times is solely due to the number of terms in the potential operator. In these calculations, where a pre-fitted, analytic potential is available, the individual energy calculations are trivial, but if this calculation were repeated using an *ab initio* electronic structure program, where each energy calculation can take anything from a few seconds to many minutes or more, the ability to reduce the number of calculations by an order of magnitude will result in a large saving of computational effort.

Figure 1(b) shows results for a second set of DD-SM simulations performed using the full databases generated by the respective full-kernel and additive kernel simulations from Fig. 1(a). The agreement between both calculations and the exact result is essentially perfect; the improvement of both fits compared to the original calculations is to be expected as a much larger region of configuration space has been used to generate the potential energy operator. In this secondary calculation, the time savings for the additive kernel over the full kernel are not as marked as in Fig. 1(a) (2,710 and 3,215 CPU seconds for full and additive kernels, respectively), because the computationally-demanding integral evaluation only occurs once, at the start of the propagation.

The conclusion of this section is that we find that the additive kernel is able to reproduce an underlying PES just as accurately as the full kernel, but requires far fewer energy evaluations; this result is very promising for grid-based quantum dynamics. However, initial attempts to combine the additive kernel method with MCTDH to allow DD-MCTDH simulations did not lead to any significant time-savings. In particular, when using the additive kernel, each database point gives rise to multiple, individual terms in the potential energy operator, one for each term in the summations in Eq. (25) (*e.g.* ten each in this case); these terms cannot be contracted down to a single operator when using MCTDH. As a result, when attempting to use the additive kernel in a DD-MCTDH scheme, the reduction in computational effort for the single-point energy evaluations and solution of the KRR linear equations is still achieved, but we do not gain anything in MCTDH propagation because there is little reduction in the number of terms in the potential energy operator.

In the next section we outline a further improvement which, in combination with the additive kernel, *is* able to reduce the number of terms in the potential energy operator, thereby allowing a significant speed up in the dynamics calculations, making DD-MCTDH calculations feasible.

IV. METHODOLOGY IMPROVEMENT II: SINGULAR VALUE DECOMPOSITION OF A GAUSSIAN PES

A. Theory and Implementation

Here, we present a new method of fitting a PES which provides another step in improving the efficiency of the DD-GB methods. Here, the KRR fitting of the PES is relegated to the secondary status of being an efficient sampling method which generates a representation of the PES; we then implement a second transformation from which another, more compact potential energy operator expression can be determined. In particular, we use the idea, demonstrated in Eq. (15), that, when using a DVR basis, we only need to know the value of the potential term at the location of the gridpoint.

To describe this second extension to our DD-GB approach, we assume that we have a KRR PES generated using the additive kernel approach described above. In the setting of our work, this KRR PES would be generated during a DD-GB simulation and the decomposition method described below can be used to accelerate evaluation of integrals over the potential energy operator (as required to accelerate MCTDH simulations). This KRR PES, referred to as $V^0(\mathbf{q})$, can then be decomposed into a simpler PES representation, appropriate for MCTDH propagation, using the following steps:

1. First, we evaluate the KRR potential energy function, V^{KRR}, at the origin of the coordinate system, $\mathbf{0}$, to set the relative, constant shift for the PES.

2. The KRR potential energy $V^{\mathrm{KRR}}(\mathbf{q})$ is evaluated at the location of the DVR gridpoints along *one* of the DOFs, with all other DOFs remaining at $q_f = 0$. This gives a one-dimensional potential energy term along

the selected mode (including any constant shift in the potential).

3. One-dimensional KRR PES slices along each of the remaining DOFs are then generated, but with the value of the KRR PES at the origin being subtracted, this time, to avoid over-counting. By this simple procedure we thus have a representation of the KRR PES as a sum of one-dimensional terms, with a single vector of numbers representing the PES along each DOF.

4. To obtain the 2-dimensional terms which couple wavepacket motion between the DOFs, we must fit the difference of the full KRR PES and the one-dimensional terms, between each unique pair of DOFs; in other words, we are fitting to the residual error between representation of the PES using only one-dimensional terms and the full KRR PES.

   To achieve this, consider two DOFs, q_g and q_h, with the coupling between them written as $V^{gh}(q_g, q_h)$. Assuming this function is separable, the integrals over the DVR basis, which appear in the wavefunction propagation, are:

$$\begin{aligned}\langle X_i^{(g)} X_j^{(h)} | V^{gh}(q_g, q_h) | X_k^{(g)} X_l^{(h)} \rangle \\ = \langle X_i^{(g)} | V^{gh}(q_g) | X_k^{(g)} \rangle \langle X_j^{(h)} | V^{gh}(q_h) | X_l^{(h)} \rangle \\ \approx V^{gh}(q_g^i) V^{gh}(q_h^j)\, \delta_{ik} \delta_{jl}\end{aligned} \tag{26}$$

   So, we find that these coupling terms, when evaluated in the DVR basis, are the product of terms on the DVR coordinate grid along the individual DOFs, g and h. We then define the residual function at any point (q_g^i, q_h^j) on this new, two-dimensional grid as

$$\begin{aligned}V^{gh}(q_g^i, q_h^j) = & V^{\mathrm{KRR}}(q_g^i, q_h^j) - V^g(q_g^i) - V^h(q_h^j) \\ & - V^{\mathrm{KRR}}(\mathbf{0})\end{aligned} \tag{27}$$

   If the DOFs, g and h, contain N_g and N_h DVR gridpoints, respectively, Eq. 27 defines an $N_g \times N_h$ matrix, $\mathbf{V}^{gh}$, the elements of which are the value of the residual at each point.

5. We can then decompose the matrix $\mathbf{V}^{gh}$ into a sum of outer products of two vectors, each term of which represents a contribution to the total coupling term. The vectors contain the values of each coupling term at the locations of the DVR gridpoints along the separate DOFs. To find such a set of vectors, we minimize the following squared Frobenius norm,

$$||\mathbf{V}^{gh} - \mathbf{V}_g^{gh} \otimes \mathbf{V}_h^{gh}||^2 \tag{28}$$

   by performing a singular value decomposition (SVD) of the coupling matrix

$$\mathbf{V}^{gh} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^T \tag{29}$$

   with the singular values in $\mathbf{\Sigma}$ being in non-decreasing order. It follows that

$$V^{gh}(q_g^i, q_h^j) = \sum_{k=1}^{\min(N_g, N_h)} \sigma_k\, u_{ik}\, w_{jk} \tag{30}$$

   We can thus set the elements of the required vectors, $\mathbf{V}_g^{gh}$ and $\mathbf{V}_h^{gh}$, which correspond to the values of the potential energy operator terms at the DVR gridpoints, to be

$$V_{g(k)}^{gh}(q_g^i) = \sqrt{\sigma_k} u_{ik} \tag{31a}$$

$$V_{h(k)}^{gh}(q_h^j) = \sqrt{\sigma_k} w_{jk} \tag{31b}$$

6. Finally, we note that, in practice, the term with the largest singular value is added, and then a new residual is created by subtracting this term from the original. If the norm of this new residual is below some pre-determined accuracy cutoff, no further terms are added; if the norm of the residual is larger than the cutoff, we loop over the other singular values in descending order, adding terms and checking the norm of the residual until convergence is reached.

7. The decomposition procedure above is repeated for all pairs of DOFs in the system.

The resulting PES representation is much more compatible with efficient MCTDH propagation because the multiple potential energy terms of the original KRR PES are contracted down to f one-dimensional terms and at most $\sum_{g<h}^{f} \min(N_g, N_h)$ two-dimensional terms. Finally, we note that this method has a similar philosophy to the POTFIT algorithm[10], which takes a PES evaluated on a grid and produces a PES in sum-of-product form by decomposing PES density matrices. An important similarity between the methods is that, in the context of grid-based dynamics methods such as MCTDH, the exact functional form of the potential is less important than obtaining the values of the potential terms at the locations of the DVR gridpoints. Of course, the underlying KRR PES used in our DD strategies is already in sum-of-products form, albeit with a sum which has too many terms for efficient simulation.

B. Salicylaldimine Proton Transfer: 4D

In this section, we demonstrate the improved computational efficiency of the SVD PES fitting method over the standard KRR approach, whilst simultaneously showing that the new method does not lead to a reduction in accuracy in the quantum dynamics.

As our benchmark problem, we again consider the 4-dimensional model of salicylaldimine proton transfer, as discussed in Section III B. We use the same computational setup with regards to DVR grid, choice of DOFs,

initial wavefunction and GPR sampling, although we focus here on the second-order additive kernel only. In addition, we also us the same fitted PES as that considered in Section III B. Wavefunction propagation was performed using MCTDH with two sets of 14 SPFs each; one describing modes ν_1 and ν_{36}, the other ν_{13} and ν_{32}. The variable mean field[10] implementation of the 6$^{\text{th}}$-order Adams-Bashforth-Moulton (ABM) integrator was used to solve the EOMs with an accuracy parameter of 10^{-5} and initial step length of 10^{-4} fs. A database of energies was created on-the-fly during propagation, and the SVD procedure, described above, was used to generate the potential energy operator for the dynamics from the KRR fit using the additive kernel. Terms from the SVD fit (Eq. (30)) were added until the Frobenius norm of the residual over all gridpoints was less than 10^{-3}. Following an initial pass in which the PES was generated on-the-fly, a second propagation was then performed using the full database of energies generated previously, with the PES fit performed at the start of the propagation and not updated further.

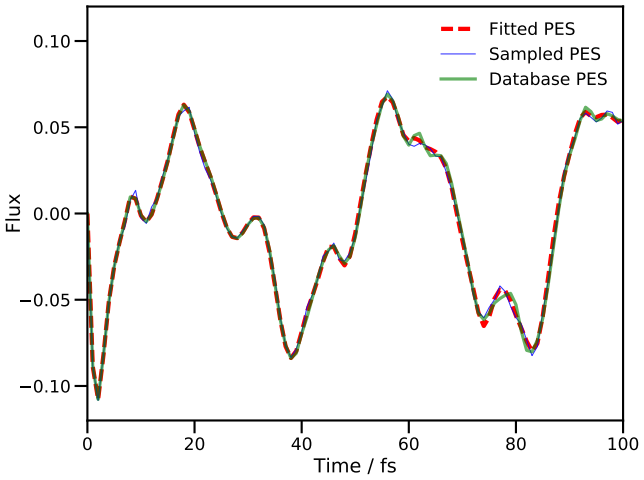

FIG. 2. Calculated proton transfer flux operator expectation value for a four DOF model of salicylaldimine. The dividing surface is placed at the barrier of the transition mode, ν_1. The dashed, red line is the result of using standard MCTDH with the fitted PES. The thin blue line is the result of a DD-MCTDH calculation with the potential energy operator constructed using SVD/additive kernel, and the thick green line uses the same SVD method, but using the PES database generated in the previous sampling calculation. All simulations used energies evaluated using the fitted PES of Polyak *et al.*[57]

In Fig. 2 we present the calculated flux expectation value along the ν_1 mode for both calculations, as well as the result from an MCTDH calculation using the same PES. There is little to say about the results in Fig. 2, except that the DD-MCTDH calculations give results in excellent agreement with the standard MCTDH calculation.

With regards to computational effort, 741 energy values were added to the PES database during the first DD-MCTDH propagation. This calculation took 3,893 CPU seconds on a single-core desktop computer, with the SVD fitting routines contributing only 59 CPU seconds to the total; in other words, the additional effort from SVD fitting is minimal. The next question is whether there is a saving in effort given by reducing the number of terms in the potential energy operator. For a one-state, four-DOF problem, there are ten potential terms for each energy value added to the database when using the standard KRR fit with the second-order, additive kernel; in the calculation reported here, this means the potential energy operator without SVD reduction would comprise 7,410 terms.

During the first calculation, when the PES is initially sampled, the number of potential energy operator terms, generated by the new SVD procedure, ranged from 39 to 56 (with a different number generated after each sampling step). The second calculation, using the pre-computed database, used a fit of 56 terms, a reduction by a factor of more than 130 over the standard KRR fit. It is also worth noting that the VCHAM fitted PES, from which we extracted the energies used in these calculations, had 40 potential operator terms (19 1-dimensional terms and 21 2-dimensional, compared to a split of 4 and 52 in the SVD fit). The SVD fit in this case requires only a few more terms than the VCHAM fit and, in fact, the DD-MCTDH calculation with the pre-computed database took less time than the standard MCTDH calculation (1,222 against 2,098 CPU seconds) due to shorter integration time-steps being taken in the latter case.

C. Salicylaldimine Proton Transfer: 6D

Having demonstrated the accuracy and reduced computational effort, we provide a further example to compare with results from our previous work[60]. In this previous case, we considered a 6-dimensional model of proton transfer in salicylaldimine, performing MCTDH dynamics using the fitted PES, and the standard KRR fit with full kernel. In that work, we found reasonable agreement between the full, MCTDH dynamics and the DD-MCTDH calculation, but with considerable computational effort. As such, we repeat these calculations here using our new DD-MCTDH scheme to further assess the benefits.

The 6-dimensional calculations included the same DOFs as used in the 4-dimensional calculations discussed above, but with addition of modes ν_{10} and ν_{11} (see Fig. (2) in reference 57). These additional modes each required 21 basis functions (harmonic oscillator DVR) and were included in an additional 2-dimensional combined mode described by 14 SPFs. The four original DOFs were treated in the same way as described in Sections III B and IV B, as was the integration of EOMs. The ini-

tial wavepacket was the product of Gaussian functions in all six DOFs, with centres and widths as described above for the four original DOFS and, additionally centred at $\langle v_{10}\rangle$=$\langle v_{11}\rangle$=0.0, with additional widths $\langle \mathrm{d}v_{10}\rangle$=0.7745 and $\langle \mathrm{d}v_{11}\rangle$=0.7590.

Figure 3 shows the flux along v_1 for three calculations: (i) standard MCTDH using the fitted PES, (ii) DD-MCTDH using the additive kernel and SVD fitting, with a database of energies from a prior calculation, and (iii) DD-MCTDH using the standard, 6-dimensional kernel with a pre-computed database of energies and using 18 SPFs per mode, as presented previously in reference 60.

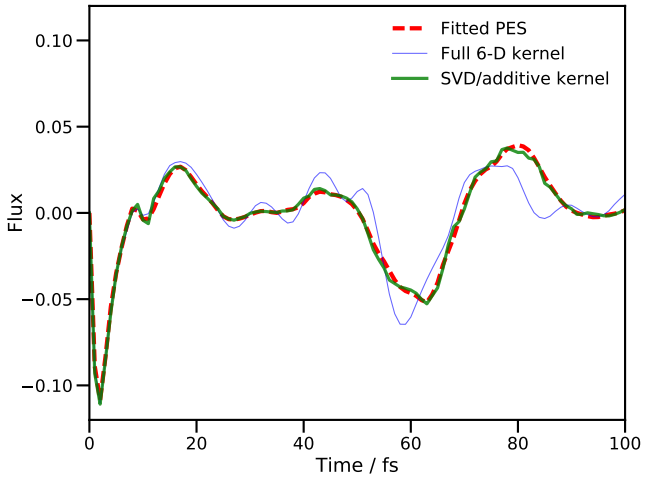

FIG. 3. Plot of the flux operator expectation value for a six degree-of-freedom model of salicylaldimine. The dividing surface is placed at the barrier of the transition mode, v_1. The solid, red line is the result of using MCTDH with the fitted PES. The hashed, blue line is the result of a DD-MCTDH calculation with the potential operator constructed using the SVD fit to an additive kernel GPR PES and the dashed, green line is the equivalent result running the calculation with the pre-calculated database. All calculations used energies calculated from the fitted PES of Polyak *et al*[57].

Figure 3 clearly demonstrates the improvement in agreement between the full MCTDH result and the DD-MCTDH calculations when using the SVD fit and additive kernel; the RMS errors in the calculated fluxes are an order of magnitude smaller when using SVD and additive kernel, compared to the full kernel calculation. We note that the effect of using 14 SPFs per combined mode as opposed to 18 in the full kernel DD-MCTDH calculation is minimal; an 18 SPF MCTDH calculation was also carried out and the plot of the flux was found to be almost identical to the corresponding 14 SPF calculation.

The errors in the full kernel case are due to the lack of convergence in the database of energies[60]; 10,101 entries were sampled, the maximum possible in that calculation, and the propagation took 143 hours on 16 processors. As a result, achieving convergence of the database would require a huge additional computational effort. In contrast, the SVD/additive kernel calculation only needed a database of 2,185 energies, once again illustrating the large saving in effort by using an additive kernel. From these 2,185 energies, a potential expansion of 134 terms was constructed by the SVD procedure, comparing very favourably with the 45,885 needed when using the additive kernel alone (six 1-dimensional, and fifteen 2-dimensional, terms *per* database entry). We also note that the VCHAM-fitted PES contained 78 terms in the six DOFs (29 1-dimensional and 49 2-dimensional), so the DD-MCTDH procedure does require more terms in the PES expansion, leading to a longer propagation (4,451 CPU seconds compared to 3,258 in the standard MCTDH calculation, both on a single processor on a desktop machine). Just over 12 hours CPU time was required for the initial propagation which generated the database for PES construction in the DD-MCTDH calculation, meaning that the full DD-MCTDH procedure with SVD and additive kernel took about 13 and a half hours; this is far less computational effort than required to perform either VCHAM fitting or a DD-MCTDH calculation using the full kernel. We also note that the use of 14 SPFs *per* combined mode as opposed to 18 for the full kernel calculation produces a saving in effort, all things being equal, but only by a factor of about 2.7 (based upon Eq. (74) in reference 10); nowhere near enough to account for the actual difference, which is mainly down to the size of the database. The use of SVD fitting with additive kernel in DD-MCTDH thus improves accuracy while simultaneously reducing computational effort.

D. Non-Adiabatic Dynamics of Pyrazine

As a final example showing the utility of the proposed SVD/additive kernel variant of the DD-MCTDH approach, we present results of a simulation modelling the non-adiabatic dynamics of pyrazine using *ab initio* electronic structure calculations. Pyrazine is a classic test case in non-adiabatic dynamics, particularly in the calculation of the absorption spectrum obtained by excitation to the S_2 excited electronic state.[15–17]. The presence of a conical intersection with the S_1 state in the vicinity of the Franck-Condon geometry makes this system an ideal test of whether a method can accurately represent the wavepacket as it moves between the electronic states.

To test our method, the geometry of pyrazine was optimized at the state-averaged complete active space self-consistent field level with 8 electrons in 8 orbitals (SA-CAS(8,8)) using the DZP basis in Molpro[79,80], the three lowest energy states being included and equally weighted. The orbitals were those in the π-system, plus the lone pair orbitals on the nitrogens. A subsequent frequency calculation generated the normal modes for use in the dynamics. A DD-MCTDH calculation was then performed

on the two excited states using the four mass-frequency scaled normal-modes v_{6_a}, v_{10_a}, v_1 and v_{9_a} (given as modes $3A_g$, $7B_{1g}$, $10A_g$ and $15A_g$ by Molpro), as used in the original MCTDH studies. Harmonic oscillator DVR basis sets were used along all four modes with 32, 22, 21 and 12 functions, respectively, One-dimensional SPFs were used for all DOFs, with different functions on each state (the multi-set formalism[10]): 7 SPFs on each state were used along v_{6_a}, with 12 on S_1 and 11 on S_2 on v_{10_a}, 6 and 5 respectively along v_1, and sets of 5 and 4 members each on v_{9_a}. The default ABM integrator was used to solve the MCTDH EOMs for 100 fs, with the initial wavefunction constructed as a product of the v=0 harmonic oscillator eigenfunctions along each mode, placed on the S_2 state and centered at the Franck-Condon point. To generate energies to construct the KRR representation of the PES, 100 points within 3 standard deviations of the wavepacket centre were sampled every femtosecond and added to the database if the variance at that geometry exceeded 10^{-3}. The energies of the two electronic states, along with the non-adiabatic couplings between them, were calculated using SA-CAS(8,8)/DZP, as described above; the energies were diabatised as outlined in Section II A prior to performing KRR fitting. Additionally, in order to maintain symmetry in the PES, each geometry was reflected in the $v_{6_a}/v_1/v_{9_a}$ plane (*i.e.* the sign of the v_{10_a} coordinate was changed and a point added at the resulting geometry). The additive kernel was used in the KRR process, and the SVD fitting procedure was used to generate the final PES on which the dynamics were performed.

As above, two calculations were performed; the first to generate the database of energies needed to represent the PES, and the second using the full database. Overall, 5,185 energies were generated during the first propagation, resulting in 335 terms in the PES expansion generated using SVD. Results from both calculations are shown in Figs. 4 and 5.

In Fig. 4, we show the absorption spectrum after vertical excitation to the S_2 state, calculated in both calculations using the Fourier transform of the autocorrelation function

$$c(t) = \langle \Psi(0) | \Psi(t) \rangle. \tag{32}$$

More details of this method can be found in reference 10, particularly regarding the use of a damping function to minimise artefacts in the spectrum.

Our calculations are not aimed at reproducing the experimental spectrum to a high degree of accuracy; the relatively low-level electronic structure method used to generate the PES precludes such efforts. We note that the original MCTDH studies of pyrazine[15–17] used a PES which was finely tuned to reproduce the experimental spectrum. Instead, our aim here is to show that a reasonably-accurate absorption spectrum can be calculated (in a reasonable wall-time) using *ab initio* calculations and our SVD/additive-kernel method. The total wall time for our first calculation was 101 hours (including all electronic structure calculations) using a single processor on a desktop machine, whilst the second propagation took just $6\frac{1}{2}$ minutes. Around 88.6% of the first MCTDH calculation time was taken up in evaluating the KRR variance (Eq. (16)); clearly there is a need for improvement here if possible.

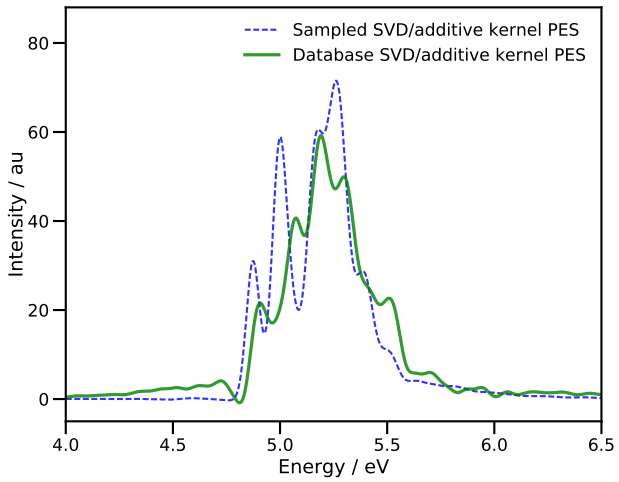

FIG. 4. Absorption spectra of pyrazine after vertical excitation to the S_2 state. Energies are relative to the S_0 minimum. The blue dashed line is the spectrum obtained in a DD-MCTDH calculation using an additive kernel and SVD fitting of the PES, with energies being added to the database on-the-fly. The green solid line is the spectrum from a DD-MCTDH calculation using the full database of energies generated during the first propagation.

The calculated spectra highlight the need to run an initial calculation (or even perhaps more than one in order to build up a converged PES) which samples configuration space before running a final calculation with a static PES database. The spectrum from the first propagation has some unphysical, negative regions (around 4.8 eV) along with tails at high and low energies, suggesting a low-accuracy representation of the PES. Furthermore, the addition of new energies to the PES database every femtosecond means that the PESs over which the wavepacket is moving are time-dependent, such that energy is not strictly conserved, leading to inaccuracies in the spectrum.

The conclusions from Fig. 4 are backed up by Fig. 5, which shows the population of the diabatic state corresponding to the S_2 adiabatic state at the Franck-Condon point. These results show significant population transfer between diabatic states, showing that our method can reproduce the diabatic couplings between states which are needed to model non-adiabatic effects such as the dynamics through conical intersections. Our calculations also show the immediate depopulation of the second excited state which was seen in the earlier MCTDH studies[15] indicating that, although the electronic structure method

used here is not highly-accurate, we are able to reproduce a key feature of the dynamics of pyrazine. The qualitative similarities in population transfer follow through the duration of the dynamics; the population reaches a minimum after 30 fs (after about 45 fs in the earlier work[15]) before a recurrence peaking after 60 fs (80 fs previously). These similarities indicate that, given further improvements in the efficiency in our algorithm, and using a more accurate electronic structure method, the DD-MCTDH method proposed here is capable of producing accurate results in a reasonable time.

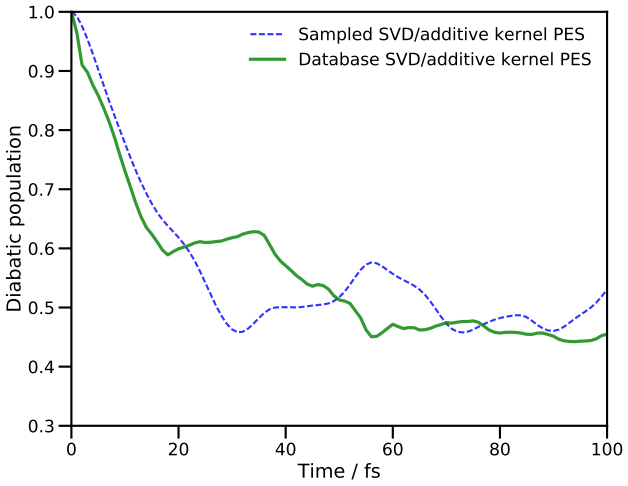

FIG. 5. Population of the second excited diabatic electronic state of pyrazine after vertical excitation. The blue dashed line is the result obtained by a DD-MCTDH calculation using an additive kernel and subsequent SVD fitting of the potential with energies being added to the database on-the-fly. The solid green line is the result of a DD-MCTDH calculation using the database of energies generated by the first propagation.

V. CONCLUSIONS

We have presented two improvements to our previously published DD method, both of which significantly improve the computational efficiency of the method: the use of an additive kernel in KRR greatly reduces the number of electronic structure calculations required to represent the PES, while the SVD fitting procedure can quickly and accurately reduce the number of terms in the resulting potential energy operator. Our simulations of proton transfer in salicylaldimine and non-adiabatic dynamics in pyrazine show that our method can be used to efficiently model quantum dynamics in many-dimensional systems.

That said, further developments are clearly warranted. As mentioned in the discussion of the pyrazine results, calculating KRR variance is a relatively expensive operation and should be addressed. Further work on the choice of the kernel is needed to determine whether an alternative would allow further reductions in the required number of electronic structure calculations. It is also possible to look into optimisation of the parameters in the kernel, particularly the coefficient of the exponent (which determines the width of the Gaussian functions); an optimal choice here will again reduce the number of electronic structure calculations needed. The use of gradient information to improve the representation of the PES is another possible avenue of investigation. We are also aware that the structure of a PES is not necessarily limited to interactions between, at most, two DOFs, and interactions between three or more DOFs may be required in an accurate expansion of the potential. Extension to include higher-order terms is ongoing, focusing on tensor decomposition to generalise the SVD fitting procedure. Finally, we note that the *a priori* choice of DOFs to include in the dynamics is an open question which we are aiming to address, particularly when dealing with larger molecules.

However, with the developments proposed here, an on-the-fly implementation of MCTDH, without the arduous task of pre-fitting a sum-of-products potential energy operator, is now a reality.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Leverhulme Trust for funding (RPG- 2016-055) and the Scientific Computing Research Technology Platform at the University of Warwick for providing computational resources. Data from Figs. 1-5 may be accessed at http://wrap.warwick.ac.uk/id/eprint/98400.

[1] N. Boekelheide, R. Salomón-Ferrer, and T. F. Miller III, Proc. Nat. Acad. Sci. USA **108**, 16159 (2011).
[2] L. Masgrau, A. Roujeinikova, L. O. Johannissen, P. Hothi, J. Basran, K. E. Ranaghan, A. J. Mulholland, M. J. Sutcliffe, N. S. Scrutton, and D. Leys, Science **312**, 237 (2006).
[3] C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández, and B. Kohler, Annu. Rev. Phys. Chem. **60**, 217 (2009).
[4] V. G. Stavros and J. R. Verlet, Annu. Rev. Phys. Chem. **67**, 211 (2016).
[5] L. A. Baker, M. D. Horbury, S. E. Greenough, F. Allais, P. S. Walsh, S. Habershon, and V. G. Stavros, J. Phys. Chem. Lett. **7**, 56 (2016).
[6] L. A. Baker, S. E. Greenough, and V. G. Stavros, J. Phys. Chem. Lett. **7**, 4655 (2016).
[7] A. H. Zewail, Science **242**, 1645 (1988).
[8] A. Douhal, S. K. Kim, and A. H. Zewail, Nature **378**, 260 (1995).
[9] A. H. Zewail, J. Phys. Chem. A **104**, 5660 (2000).
[10] M. H. Beck, A. Jäckle, G. A. Worth, and H. D. Meyer, Phys. Rep. **324**, 1 (2000).
[11] M. Schröder, F. Gatti, and H.-D. Meyer, J. Chem. Phys. **134**, 234307/1 (2011).
[12] M. Schröder and H.-D. Meyer, J. Chem. Phys. **141**, 034116/1 (2014).
[13] M. Coutino-Neto, A. Viel, and U. Manthe, J. Chem. Phys. **121**, 9207 (2004).
[14] K. Sadri, D. Lauvergnat, F. Gatti, and H.-D. Meyer, J. Chem. Phys. **141**, 114101/1 (2014).

[15]G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **105**, 4412 (1996).
[16]G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **109**, 3518 (1998).
[17]A. Raab, G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **110**, 936 (1999).
[18]S. Neville and G. Worth, J. Chem. Phys. **140**, 034317/1 (2014).
[19]T. J. Penfold and G. A. Worth, J. Chem. Phys. **131**, 064303/1 (2009).
[20]M. D. H. Köppel, I. Baldea, H.-D. Meyer, and P. Szalay, J. Chem. Phys. **117**, 2657 (2002).
[21]H. Wang and M. Thoss, J. Chem. Phys. **119**, 1289 (2003).
[22]H. Wang, J. Phys. Chem. A **119**, 7951 (2015).
[23]O. Vendrell and H.-D. Meyer, J. Chem. Phys. **134**, 044135/1 (2011).
[24]T. Hammer and U. Manthe, J. Chem. Phys. **134**, 224305/1 (2011).
[25]K. Giri, E. Chapman, C. S. Sanz, and G. Worth, J. Chem. Phys. **135**, 044311/1 (2011).
[26]H. Köppel, W. Domcke, and L. S. Cederbaum, Adv. Chem. Phys. **57**, 59 (1984).
[27]L. S. Cederbaum, H. Köppel, and W. Domcke, Int. J. Quant. Chem. **15**, 251 (1981).
[28]H.-D. Meyer, F. Gatti, and G. A. Worth, eds., *Multidimensional quantum dynamics: MCTDH theory and applications* (Wiley, Weinheim, Germany, 2009).
[29]M. Persico and G. Granucci, Theo. Chem. Acc. **133**, 1526/1 (2014).
[30]G. A. Worth, M. A. Robb, and B. Lasorne, Mol. Phys. **106**, 2077 (2008).
[31]J. C. Tully and R. K. Preston, J. Chem. Phys. **55**, 562 (1971).
[32]J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).
[33]M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, J. Phys. Chem. Lett. **3**, 3090 (2012).
[34]M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, J. Chem. Theory Comput. **7**, 1253 (2011).
[35]J. C. Tully, Farad. Discuss. **110**, 407 (1998).
[36]D. F. Coker, in *Computer Simulation in Chemical Physics*, edited by M. P. Allen and D. J. Tildesley (Kluwer Academic, Dordrecht, 1993) pp. 315–377.
[37]M. S. Topaler, M. D. Hack, T. C. Allison, Y.-P. Liu, S. L. Mielke, D. W. Schwenke, and D. G. Truhlar, J. Chem. Phys. **106**, 8699 (1997).
[38]B. R. Smith, M. J. Bearpark, M. A. Robb, F. Bernardi, and M. Olivucci, Chem. Phys. Lett. **242**, 27 (1995).
[39]M. J. Bearpark, F. Bernardi, M. Olivucci, M. A. Robb, and B. R. Smith, J. Am. Chem. Soc. **118**, 5254 (1996).
[40]R. Mitrić, J. Petersen, and V. Bonacić-Koutecký, Phys. Rev. A **79**, 053416/1 (2009).
[41]P. Lisinetskaya and R. Mitrić, Phys. Rev. A **83**, 033408/1 (2011).
[42]M. D. Hack, A. Jasper, Y. L. Volobuev, D. W. Schwenke, and D. G. Truhlar, J. Phys. Chem. A **103**, 6309 (1999).
[43]M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).
[44]M. Ben-Nun and T. J. Martínez, Adv. Chem. Phys. **121**, 439 (2002).
[45]H. R. Hudock, B. G. Levine, A. L. Thompson, H. Satzger, D. Townsend, N. Gador, S. Ullrich, A. Stolow, and T. J. Martínez, J. Phys. Chem. A **111**, 8500 (2007).
[46]T. Mori, W. Glover, M. Schuurman, and T. Martínez, J. Phys. Chem. A **116**, 2808 (2012).
[47]G. Worth and I. Burghardt, Chem. Phys. Lett. **368**, 502 (2003).
[48]I. Burghardt, H.-D. Meyer, and L. S. Cederbaum, J. Chem. Phys. **111**, 2927 (1999).
[49]G. Richings, I. Polyak, K. Spinlove, G. Worth, I. Burghardt, and B. Lasorne, Int. Rev. Phys. Chem. **34**, 269 (2015).
[50]G. A. Worth, M. A. Robb, and I. Burghardt, Farad. Discuss. **127**, 307 (2004).
[51]B. Lasorne, M. J. Bearpark, M. A. Robb, and G. A. Worth, J. Phys. Chem. A **112**, 13017 (2008).
[52]B. Lasorne, M. A. Robb, and G. A. Worth, Phys. Chem. Chem. Phys. **9**, 3210 (2007).
[53]D. Mendive-Tapia, B. Lasorne, G. Worth, M. Bearpark, and M. Robb, Phys. Chem. Chem. Phys. **12**, 15725 (2010).
[54]C. Allan, B. Lasorne, G. Worth, and M. Robb, J. Phys. Chem. A **114**, 8713 (2010).
[55]D. Asturiol, B. Lasorne, G. Worth, M. Bearpark, and M. Robb, Phys. Chem. Chem. Phys. **12**, 4949 (2010).
[56]M. Araújo, B. Lasorne, A. Magalhaes, M. Bearpark, and M. Robb, J. Phys. Chem. A **114**, 12016 (2010).
[57]I. Polyak, C. Allan, and G. Worth, J. Chem. Phys. **143**, 084121/1 (2015).
[58]M. Vacher, M. J. Bearpark, and M. A. Robb, Theor. Chem. Acc. **135**, 187/1 (2016).
[59]G. Richings and S. Habershon, Chem. Phys. Lett. **683**, 228 (2017).
[60]G. Richings and S. Habershon, J. Chem. Theory Comput. **13**, 4012 (2017).
[61]C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, Cambridge, Massachusetts, 2006).
[62]C. Williams, in *Handbook of Brain Theory and Neural Networks*, edited by M. Arbib (The MIT Press, Cambridge, Massachusetts, 2002) pp. 466–470.
[63]A. Bartók and G. Csányi, Int. J. Quantum. Chem. **115**, 1051 (2015).
[64]J. P. Alborzpour, D. P. Tew, and S. Habershon, J. Chem. Phys. **145**, 174112/1 (2016).
[65]L. Mones, N. Bernstein, and G. Csányi, J. Chem. Theory Comput. **12**, 5100 (2016).
[66]J. Quinonero-Candela, C. E. Ramussen, and C. K. I. Williams, in *Large-Scale Kernel Machines*, edited by L. Bottou, O. Chapelle, D. DeCoste, and J. Weston (The MIT Press, Cambridge, Massachusetts, 2007) pp. 203–224.
[67]D. Duvenaud, H. Nickisch, and C. E. Rasmussen, in *Neural Information Processing Systems Conference* (2011).
[68]K. Chalupka, C. K. Williams, and I. Murray, J. Mach. Learn Res. **14**, 333 (2013).
[69]P. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).
[70]J. Frenkel, *Wave Mechanics* (Clarendon Press, Oxford, 1934).
[71]G. Worth, K. Giri, G. Richings, I. Burghardt, M. Beck, A. Jäckle, and H.-D. Meyer, "The Quantics Package, Version 1.1," Tech. Rep. (University of Birmingham, Birmingham, UK, 2016).
[72]G. Richings and G. Worth, J. Phys. Chem. A **119**, 12457 (2015).
[73]G. Richings and G. Worth, Chem. Phys. Lett. **683**, 606 (2017).
[74]B. Esry and H. Sadeghpour, Phys. Rev. A **68**, 042706/1 (2003).
[75]M. Baer, Chem. Phys. Lett **35**, 112 (1975).
[76]H. Rabitz and O. Aliş, J. Math. Chem. **25**, 197 (1999).
[77]O. Aliş and H. Rabitz, J. Math. Chem. **29**, 127 (2001).
[78]T.-S. Ho and H. Rabitz, J. Chem. Phys. **119**, 6433 (2003).
[79]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, "Molpro, version 2015.1, a package of ab initio programs," (2015).
[80]H.-J. Werner and P. J. Knowles, J. Chem. Phys. **82**, 5053 (1985).